

MEMORANDUM
RM-5868-DOT/RC
APRIL 1969

PREFERENCES FOR
MULTI-ATTRIBUTED ALTERNATIVES

Howard Raiffa

PREPARED FOR:
U. S. DEPARTMENT OF TRANSPORTATION
FEDERAL RAILROAD ADMINISTRATION
OFFICE OF HIGH SPEED GROUND TRANSPORTATION

The RAND *Corporation*
SANTA MONICA • CALIFORNIA

MEMORANDUM
RM-5868-DOT/RC
APRIL 1969

PREFERENCES FOR MULTI-ATTRIBUTED ALTERNATIVES

Howard Raiffa

This research is supported by the U.S. Department of Transportation, Federal Railroad Administration, Office of High Speed Ground Transportation under Contract 3-0008, and by The Rand Corporation as part of its self-sponsored research in the public interest.

Copyright © 1969
THE RAND CORPORATION

This study is presented as a competent treatment of the subject, worthy of publication. The Rand Corporation vouches for the quality of the research, without necessarily endorsing the opinions and conclusions of the authors.

Published by The RAND Corporation

PREFACE

This Memorandum is one in a series being prepared by The RAND Corporation for the Northeast Corridor Transportation Project, Office of High Speed Ground Transportation, U.S. Department of Transportation. The overall research effort is directed toward development of comprehensive and systematic methodology for evaluating the potential utility of alternative transportation proposals.

The series of Memoranda can be classified into several types of papers. In the first and major part we are attempting to integrate our efforts into an overall evaluation framework.* Part II is composed of supporting Memoranda on some of the more relevant theoretical aspects associated with combining many dimensions in an alternative selection process. Part III consists of important gap-filling papers and background materials.

This Memorandum is one of three thus far dealing with techniques for analysis of multi-dimensional alternatives.** The paper discusses some techniques a decisionmaker might employ if he wants to assess a utility function (interpreted in a probabilistic sense) over complex consequences--consequences that can be at best described only in terms of several attributes or descriptors. Portions of the paper extend the discussion found on pages 246-255 of the author's book, Decision Analysis: Introductory Lectures on Choices Under Uncertainty, Addison-Wesley (1968). Unlike other Memoranda in the series, this research was jointly supported by the Department of Transportation and The RAND Corporation's own research funds. The Memorandum is part of a continuing program of study and therefore findings should be considered as interim in nature.

* This material will be found in RM-5869-DOT, Measurement and Evaluation of Transportation System Effectiveness, by Frederick S. Pardee, et al, The RAND Corporation, 1969 (forthcoming).

** The other two Memoranda are RM-5865-DOT, Assessing Alternative Transportation Systems by James R. Miller III, The RAND Corporation, April 1969, and RM-5877-DOT, Improving the System Design and Evaluation Process by the Use of Trade-off Information: An Application to Northeast Corridor Transportation Planning by K. R. MacCrimmon, The RAND Corporation, April 1969.

SUMMARY

This paper presents several techniques for assessing the utility of complex alternatives. It studies the problem where each alternative can be described by a series of attributes which are interdependent in various ways. A hierarchical structuring procedure for obtaining a relevant list of attributes is described and examples are given from Northeast Corridor transportation problems and from medical treatment problems.

The analysis begins by assuming that each alternative or consequence can be described by a sequence of numbers or numeraires, with the i th number interpreted as its score or performance on the i th attribute. The basic problem is: How does one assess a utility function over an r -tuple of numeraires?

Certain techniques are presented for reducing the case $r = 2$ to the case $r = 1$. The use of constant or variable substitution rates is basic here. The discussion of the general case divides into two parts: additive and nonadditive representations. The issue is whether or not one can assess the utility of a complex consequence by assessing utilities on each dimension or component and then forming a weighted average. Critical in any such discussion is a precise notion of independence of attributes, and two such notions are introduced: weak conditional utility independence and strong conditional utility independence. Using these ideas, some of the basic results on additive representations are recalled and reformulated, including those of Miller and Fishburn.

Turning to nonadditive representations, an area quite neglected in the past, the paper presents sufficient conditions in terms of strong conditional utility independence for obtaining a quasi-additive representation.

Finally, some of the ideas developed here are applied to various problems. In particular, they are applied to the problem of assessing the "value of a life," an issue of importance in both transportation and medical treatment applications. And they are applied to the problem of bringing in intertemporal tradeoffs (one might be willing to give up something now to get more later) in analyzing utilities.

ACKNOWLEDGEMENTS

The first draft of this paper was written at RAND in July 1968. For this draft the author used some very rough, condensed lecture notes prepared the previous spring for a course (Economics 218: Decision Analysis) given at Harvard.

The author would like to express his appreciation to F. C. Iklé for inviting him to RAND for the summer to work on this important topic, to E. S. Quade for expressing a continuing interest in the work, to F. S. Pardee for stimulating its specific relationship to problems in the field of transportation system analysis, to F. Roberts and L. J. Savage for their careful review of the manuscript, and to many others at RAND with whom the author engaged in extensive and profitable dialogue.

The present draft profits from constructive feedback received from members of the Harvard Business School Decision Process Seminar, chaired by Raymond Bauer. In addition the author is indebted to Ralph Keeney, James Campen and Bing Sung for their helpful comments.

CONTENTS

PREFACE	iii
SUMMARY	v
ACKNOWLEDGEMENT	vii
<hr/>	
Section	
1. Introduction	1
2. Preliminary Formulation	3
3. Representation of C's in Terms of Numeraires	5
1. Single Numeraire	5
2. Several Numeraire Case	6
4. Techniques for $r = 2$	8
1. Constant Substitution	8
2. Variable Substitution	12
3. A Generalization of Variable Substitution	15
4. General Remarks	16
5. Hierarchical Structures of Attributes	17
1. Introduction	17
2. Goals, Specifications, Means, and Ends	18
3. Conditional Preferences	24
4. Utility Independence and Decomposition	26
5. Examples from Medical Treatment	29
6. Additive Representations	35
1. Acknowledgments	35
2. Case of Two Summands	35
3. Generalization to More Summands	37
4. Relative Scaling	39
5. An Example	43
6. Miller's Additive Representation of Worth	50
7. Use of a Classical, Additive-Utility Theorem	53
8. One More Proof of the Justification of Subjective Probabilities	63
7. Non-Additive Representations	66
1. Quasi-Additive Representations	66
2. Quasi-Additive and Additive Representations	71
3. Use of Certainty Equivalents	74
4. Asymmetric Forms of Utility Independence	75
8. Value of a Life	81
1. Value of YOUR Life	81
2. Value of a Statistical Life	93

9. Intertemporal Tradeoffs	95
1. Introduction	95
2. The So-Called Certainty Case	96
3. Time Resolution of Uncertainty	98
Bibliography	107

PREFERENCES FOR MULTI-ATTRIBUTED CONSEQUENCES

1. Introduction

In this paper I shall consider some techniques for grappling with a problem that is particularly vexing and seems to me at least to be at the heart of the methodological difficulties in most systematic analyses of complex problems. In rough terms the problem is: How can a decision maker (say YOU, or myself) think systematically about choosing the "best" of several alternative actions when each action will result in one of several possible complex outcomes and where each outcome can at best be described only in terms of its performance characteristics on many diverse attributes? For example:

In a business context the business executive may be concerned with his corporation's earnings stream over time, with stock price, with share of the market, with goodwill, with labor relations, with fulfilling his obligations to society, with his "images" to different groups, and so on.

In a government context the administrator in reviewing a proposed course of action might be concerned with the time stream of expenditures in terms of capital outlays, of allocation of managerial talent, and of the time stream of benefits and the incidence of these benefits to different sectors. These benefits may themselves be a composite bundle of incommensurable entities.

In a medical context any one specific outcome that might result when a patient undergoes a given medical treatment may be described in terms of: the cost of the treatment, days in bed with extreme discomfort, with medium discomfort, and with mild discomfort, days taken for recuperation, probability

of a relapse, probability of awesome complication A, complication B, and so forth. And in some cases the patient's doctor -- the decision maker in this case -- may have to worry about the implications of an outcome to himself, to members of the patient's family, and to society in general.

In this paper I shall emphasize the aspect of the problem dealing with "values", with the problem of handling the "incommensurables" -- and I shall not stress the uncertainty aspect of the problem. But as I see it, it is important to keep the uncertainty aspect in mind and to attempt to develop a methodology for the value problem that could incorporate an uncertainty analysis as well.

My approach can be termed conditionally prescriptive: . . . If you believe these ways in these simple problems and if you want to act in a manner consistent with these rules, then you should behave thus in these complex problems. In this paper I will completely slight the descriptive approach: How decisions are in fact made.

Much has been written about choice problems of the kind I will discuss and while I could spend time saying what I think is wrong with a given approach I would rather confine my remarks to the elaboration of techniques that I think are promising, both for theory and practice.

Before one gets to the stage of evaluating alternatives in a given problem domain, one must first recognize that a problem exists and one should seek imaginative alternatives for review. I will not deal with these obviously important facets of methodology. My concern is with choice amongst previously specified alternatives.

I shall also make the assumption that there is a single decision maker. If the decision making unit comprises a group of individuals, I will not consider techniques that the group members could use to negotiate a group consensus.

2. Preliminary Formulation

An abstract, simplified version of the problem I have in mind can be posed as follows: A decision maker ("DM" henceforth) has to choose between two acts: a_1 and a_2 . If he chooses act a_i (for $i = 1, 2$) and event E_j (for $j = 1, 2, \dots, J$) occurs, then a consequence (or outcome) C_{ij} will result. See Figure 1. Since I do not want to emphasize the probabilistic

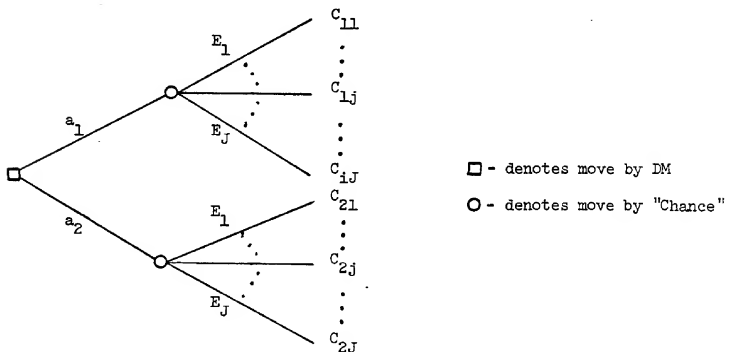


Figure 1

aspect of the problem let us assume that event E_j has an associated objective, frequency-based probability assignment, $P(E_j)$, that is known to all concerned. Following the school of thought that believes in the maximization of expected utility, our task is to assign a utility number $u(C_{ij})$ to consequence C_{ij} and to compute the two expected utility indices;

$$\sum_{j=1}^J P(E_j)u(C_{1j}) \quad \text{and} \quad \sum_{j=1}^J P(E_j)u(C_{2j}),$$

which purportedly reflect the relative desirability of act a_1 versus a_2 ; the higher the expected utility index, the better the associated act. In this paper I shall be concerned mainly with the association of utility numbers to consequences when these consequences are complicated stimuli. This assessment of utility values is far from easy but there are many tricks of the trade that can help systematize one's thought process and in some circumstances crude evaluations might suffice to resolve the action problem; but if these crude evaluations do not suffice, then at least we will know how to refine the evaluations.

The problem posed in Figure 1 seems a bit too special, but really it isn't. It is a trivial matter to add more acts, to let the ensuing potential events that flow from an act to be dependent on that act -- that is, for act a_1 we could consider events $E_1^{(i)}$, $E_2^{(i)}$, Furthermore, if we can solve the static problem as posed in Figure 1, then it is easy to extend the results to a more dynamic approach (e.g., to incorporate decision trees), which would allow us to treat such topics as the optional accumulation of evidence and information, flexibility of action over time, and so on.

Now let's turn to our main topic: given a complex consequence C how can we go about assigning a utility number to C , a number that mirrors our "liking" for C and is appropriate to use in the expected value operation. If we leave aside for the moment questions of practicality in measurement, then we could proceed as follows:

(1) choose two reference consequences, C_* and C^* say, such that any consequences under consideration is at least as good as C_* and at most as good as C^* ;

(2) for any consequence C , let the DM determine a number $u(C)$ between 0 and 1 such that the DM is indifferent between getting C for certain and getting a lottery which yields C^* with (objective) probability $u(C)$ and C_* with the complementary probability $[1 - u(C)]$.

By means of the above procedure we obtain a function u that associates numbers to consequences and, in theory, these numbers satisfy all the requirements we demand of them, namely: it is appropriate in probabilistic choice situations to maximize the expected value of these u -values. But this is a wholly impractical way to determine the u -function in complicated problems. The real problem is: faced with an important decision problem for which we want to make a critical, detailed, and systematic analysis, can we, in practice, assess utility values for complex consequences in a responsible manner?

3. Representation of C's in Terms of Numeraires

3.1. Single Numeraire

In textbook applications of utility theory one usually considers the case where each consequence is described in terms of a single numeraire such as an

incremental cash flow or monetary asset position. If x represents the generic value of this numeraire, then the utility problem boils down to defining a function u over this real variable. But now one does not have to find for each and every x a utility value $u(x)$ from basic principles; rather one can exploit the natural preference ordering for this numeraire -- for example, more x is preferred to less and therefore the DM knows at the outset he wants u to be monotonic. Again from general qualitative considerations he might know a great deal more: he might want his u -function to be monotonic, to be continuous, to be concave in shape if he is risk-averse (i.e., if for any x he prefers x for certain to a fifty-fifty chance at $x - h$ and $x + h$ for any positive h), and to have more subtle properties like decreasing risk-aversion (i.e., to have $-u''(x)/u'(x)$ be monotonically decreasing in x). It is qualitative considerations of this kind that constrain the u -function sufficiently so that it becomes possible to determine a suitable function with operational significance. But our real problem in this paper is not this well-behaved case where consequences have already been scaled in terms of a single meaningful numeraire. In passing it is desirable to emphasize once again that the underlying numeraire does not have to be money. It could be number of cures, or hours worked, or millions of MCF's of gas reserves, ..., or any other single index which captures the essence of the consequences under consideration.

3.2. Several Numeraire Case

In most applications it is not easy to summarize the essentials of the consequence by means of a single numeraire. But it might be possible to associate to each consequence C a sequence of numbers $[x_1(C), x_2(C), \dots, x_r(C)]$ which, for all practical purposes, sufficiently summarizes all the relevant

information in C . The number $x_i(C)$ can be interpreted as the index, level, or score of C on the i th criterion or attribute. In a business context attribute 1 may be a cash amount, attribute 2 a share of the market, attribute 3 an index of goodwill, and so forth. In a medical context attribute 1 may be cost of treatment, attribute 2 days of extreme discomfort, attribute 3 days for recuperation with bed rest, attribute 4 the probability of getting a relapse (after the cutoff-date of the analysis), and so forth. In other contexts $x_i(C)$ may refer to individual i 's benefits and a benevolent dictator (administrator) may wish to consider these benefits to others in his own utility function. In still another context $x_i(C)$ might denote the net cash flow in period i and then the set of numbers would depict the cash flow -- or perhaps consumption flow -- over time. Or perhaps $x_i(C)$ might denote a value index to the society of a given policy if an uncertain event E_i does in fact occur by a given date.

We now wish to develop some techniques to help a DM think systematically about assessing a utility function over an r -tuple of numeraires (x_1, x_2, \dots, x_r) . Keep in mind that if \underline{x}' and \underline{x}'' are two such r -tuples (hereafter called "points"), then the utility function u not only has to reflect ordinal preferences (i.e., the better the r -tuple the higher its associated utility number will be) but also the utility of the lottery giving p_i chance at $\underline{x}^{(i)}$ (where $\underline{x}^{(i)}$ is short for $\{x_1(C_i), \dots, x_r(C_i)\}$), for $i = 1, 2, \dots, n$, must be

$$p_1 u(\underline{x}^{(1)}) + p_2 u(\underline{x}^{(2)}) + \dots + p_n u(\underline{x}^{(n)}) .$$

The dimensionality of \underline{x} (i.e., the magnitude of r) may in some instances

be very large. Indeed in one study* I have seen the dimensionality is in the thousands. Naturally any procedure that attempts to cope with this order of complexity must somehow structure the attributes in some hierarchical arrangement. I propose to start this discussion in Section 5 and at that time we will consider such topics as goals, objectives, means, ends, and the selection of attributes for consideration (after all, these are not unique!). So far, nowhere have I said that attribute i is "independent" (in some sense) from attribute j ; indeed j might only be a minor modification of i .

Before turning our attention to the development of techniques for large r let us first consider the case where $r = 2$. Later, when we deal with r values larger than 2, we will often exploit some of the simple tricks used for $r = 2$.

4. Techniques for $r = 2$

To avoid subscripts in this section, let the typical point be (x, y) instead of (x_1, x_2) . In the (x, y) plane, let us also assume, for concreteness' sake, that preferences are to the northeast: i.e., both x and y represent values of desirable commodities.

4.1. Constant Substitution

Suppose that the iso-preference curves in the x, y plane (or in restricted portion of the x, y plane that includes all the points of interest) are parallel straight lines as in Figure 2. In this case there is some substitution rate λ , say, such that the DM is indifferent to a decrement of a unit of y as long as there is a compensating increment of λ units of x . Hence the point (x, y) will

* "Methodologies for analyzing the comparative effectiveness and costs of alternative space plans," by Dole et al, RM-5696-NASA, RAND Corp., August, 1968.

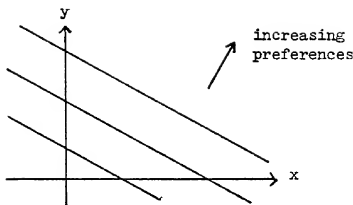


Figure 2

be indifferent to the point $(x + \lambda y, 0)$. [Alternatively we would say that (x, y) is indifferent to $(0, y + x/\lambda)$.] Hence we can summarize the (x, y) point by the single numeraire $x + \lambda y$.

Now consider a lottery where with probability p_i one obtains C_i , and assume that we can sufficiently summarize C_i by the pair (x_i, y_i) ; by substitution replace (x_i, y_i) by $(x_i + \lambda y_i, 0)$; by convention agree to abbreviate $(x_i + \lambda y_i, 0)$ by $x_i + \lambda y_i$; and finally assess a utility index u_i for the single numeraire $x_i + \lambda y_i$, keeping in mind that the second component is placed at 0. This procedure is depicted in Figure 3. The symbol $A \sim B$ is read "A is indifferent to B" which is elliptical for "the DM is indifferent between A and B". Throughout I will use the Principle of Substitution which in rough form states that the desirability of a lottery

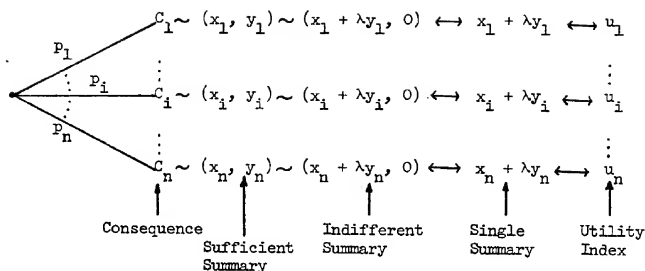


Figure 3

is not changed when any given consequence is replaced by an indifferent consequence.

As a simple illustration let (x, y) be the amount of cash flows in period 1 and 2 respectively. Since money in period 1 can be reinvested to produce more money tomorrow, or because consumption today may be sweeter than consumption tomorrow, the decision maker might decide that he is indifferent between (x, y) and $(x + \lambda y, 0)$ or $(0, y + x/\lambda)$. In this case we can think of λ as the (subjective) effective discount rate. It is then a matter of convenience whether one chooses as numeraire discounted values $x + \lambda y$ or accumulated values $y + x/\lambda$.

In both these situations it seemed natural to push one of the two components to zero. However, it may be more meaningful in some circumstances to choose a value y^* , say, and use the result that

(x_i, y_i) is indifferent to (x_i^*, y^*) , all i ,
where

$$x_i^* = x_i + \lambda(y_i - y^*) .$$

Be sure to note the omission of the subscript on y^* . Now when one considers putting a utility value on the scalar numeraire x_i^* (for $i = 1, \dots, n$), one must keep in mind that the second component is at the constant value y^* for each of the n consequences.

Some other examples where constant substitution rates may be appropriate are:

- In treating a given illness with different drugs, the x -component might represent the proportion of cures and the y -component the proportion of patients having given side-effects. Every additional increment Δ in side effects might be deemed indifferent to a $\lambda\Delta$ decrement in cures; or symbolically, one might be able to choose λ such that

$$(x, y) \sim (x - \lambda\Delta, y + \Delta)$$

for all x, y, Δ values.

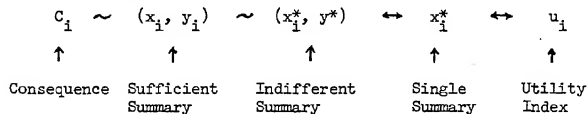
- In traffic safety (or perhaps in military matters) different strategic alternatives might result in differing numbers of fatalities (x) and casualties (y). For some analyses it might suffice to have a constant tradeoff between

Δ more casualties and ΔΔ less fatalities.

Later in the paper we will give examples where there is no appropriate constant substitution rate between x and y but where there is a constant substitution rate between a transformation of x and a transformation of y.

4.2. Variable Substitution

In general iso-preference curves in the (x, y) plane are not parallel straight lines. The substitution rate between a unit of y and x will, in general, depend on the levels of both x and y. If we have the indifference map -- this is a big If -- we can proceed almost as before. We could, for example, choose a value y^* , say, and, for each (x_i, y_i) pair find the point on the same indifference curve of the form (x_i^*, y^*) ; then we could proceed to consider the single numeraire x_i^* , and finally, keeping the value of y^* in mind, associate a utility to each of the x_i^* 's. Thus, using the symbolism of Figure 3, we would have the chain:



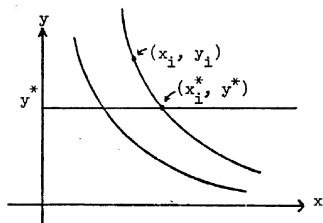


Figure 4

We may choose not to get the full indifference map. Indeed, all we want to get is the point (x_i^*, y^*) which is indifferent (subjectively) to the point (x_i, y_i) . A judicious choice of y^* will often help these subjective indifference assessments. Sometimes it is easiest to choose y^* as the minimum or maximum of the y_i 's; other times the y^* may be chosen as some natural focal point near the median of the y_i 's.

Of course, we could have chosen an x^* instead of y^* and scaled (x_i, y_i) into (x^*, y_i^*) . One must be imaginative in choosing the most convenient reduction procedure. In some circumstances, for example, it may be natural to expect y to be approximately a multiple γ of x . In this case we might choose to scale (x_i, y_i) in terms of $(x_i^*, \gamma x_i^*)$.

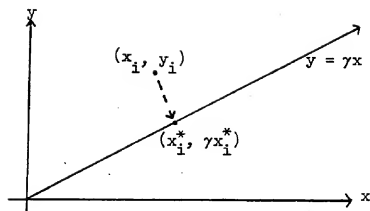


Figure 5

Example: (a) Consider once again the case where x and y are respectively cash flows in two successive periods. We might have in mind re-investing all of x into another project which will produce an amount $\Delta y = h(x)$. It may be that for small x all we can get is $(1 + i)x$ where i is a standard lending rate but for larger x we have a very profitable investment in mind; for still larger x 's we have other investments in mind that are not so very profitable, and so on. The function $h(\cdot)$ might be as depicted in Figure 6. In this case we get that (x, y) is indifferent to $[0, y + h(x)]$.

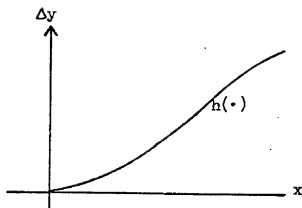


Figure 6

(b) Let x_i and y_i represent respectively consumption in periods 1 and 2. Let "base states" of consumption be C_1^* and C_2^* in periods 1 and 2. Let the decision maker decide on a value η_i , say, such that (x_i, y_i) is indifferent to $(\eta_i C_1^*, \eta_i C_2^*)$. The decision maker may find it convenient to assess η_i for each i and then to place a utility function over the η -numeraire.

4.3. A Generalization of Variable Substitution

The technique described in this section can be easily and most helpfully generalized. For example, a firm may be primarily interested in profitability, but there may be a host of other considerations which, more or less, may have to be considered, such as: orderly labor relations, public image to stockholders and to peer groups, share of the market, non-monetary selfish interests of the management elite, and so forth. It might be possible to choose a "base state", \underline{v}^* , which describes a "typical" profile of these other second-order characteristics. Now we

can think of consequence C_1 as giving rise to an evaluation (x_1, y_1) where y_1 is a prose description of everything but non-monetary direct profits. The trick of the game is now to find that value x_1^* such that (x_1, y_1) is indifferent to (x_1^*, y^*) . If y_1 is less desirable than y^* then we are asking, "considering you are at (x_1, y_1) , how much would you just be willing to change x_1 in order to go from y_1 to the "base state" y^* ?" I hope you don't infer from this that I believe this question should be answered by a snap judgment. It may take considerable soul searching and the analysis might involve a good deal of data gathering and number pushing. However, once we scale (x_1, y_1) in terms of (x_1^*, y_1^*) we can proceed as before to transform x_1^* into utility units -- keeping in mind that all second-order considerations have been adjusted to base state, y^* .

4.4. General Remarks

a. There is a general procedure, called lexicographical ordering, that is somewhat similar to the procedure discussed above, and goes as follows: Suppose that a DM must make a paired comparison between (x_1, y_1) and (x_2, y_2) , where the first attribute is considered much more important than the second; it is suggested that he first compare the x-component: if $x_1 > x_2$, then the pair (x_1, y_1) is deemed better than (x_2, y_2) -- written $(x_1, y_1) > (x_2, y_2)$; if $x_2 > x_1$, then $(x_2, y_2) > (x_1, y_1)$; if $x_1 = x_2$, then and only then does the y-component become operative and the ranking is determined by y_1 versus y_2 . In other words, the less important component gets into the act only if there are ties in the more important component. The lexicographical procedure is "objective" in the sense that

it does not require subjective substitution rates, but, in my opinion, the quest for objectivity in these circumstances is most often inappropriate since the problem simply demands subjective evaluations.

b. Another common procedure sets up an aspiration level for the y-component, say y^* , and ranks (x_i, y_i) pairs according to the magnitude of the x-component provided that $y_i \geq y^*$. Let the maximum of the x_i 's, subject to the condition that $y_i \geq y^*$, be $M_x(y^*)$. By examining how M_x behaves as y^* is varied, the analyst can subjectively choose an appropriate value for y^* . This technique has some desirable properties in choice amongst certainties but it presents difficulties when probabilistic considerations are brought into the picture, especially when the x and y components represent dependent random variables.

5. Hierarchical Structures of Attributes

5.1. Introduction

This section draws heavily on the papers by Miller [14] and by Mannheim and Hall [13]. Both papers present methods for choosing amongst alternatives in complex problems involving many objectively-incommensurable attributes. Both papers, written quite independently of each other, stress the hierarchical arrangements of attributes in terms of primary, secondary, tertiary, ..., goals.

Miller's paper assigns to each complex consequence a single "worth" or "gratify" value that is arrived at by an additive process. In contrast to utility values, however, these Miller-values are not necessarily appropriate numbers to manipulate in probabilistic situations (e.g., to use in expected value operations).

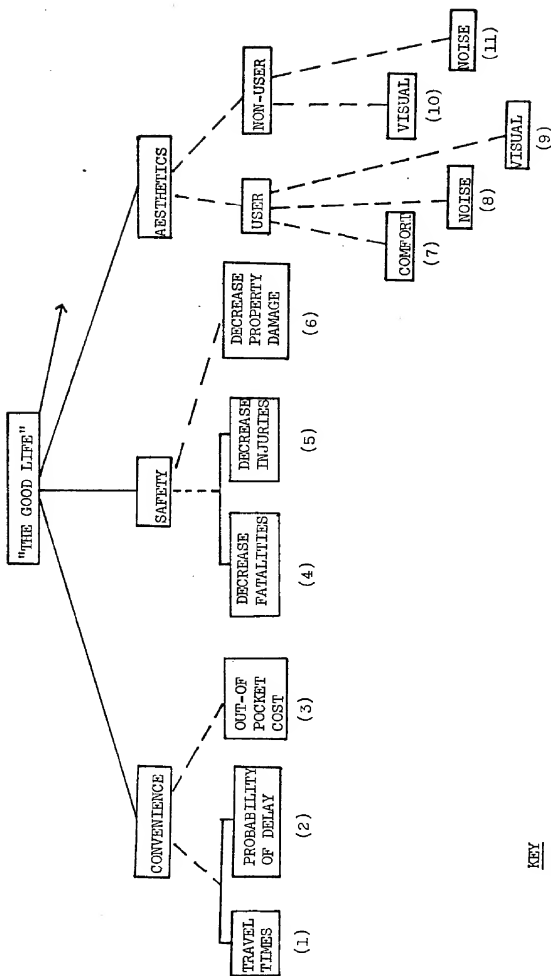
The Manheim-Hall paper rejects the utility approach as being too unworkable and the paper merely attempts to sketch out the so-called goal-fabric of the problem in qualitative terms. In some circumstances this qualitative analysis taken together with ordinal preferences is all that is necessary to choose a best action. But a lot depends on lucky dominance arguments.

5.2. Goals, Specifications, Means, and Ends

We follow Manheim-Hall [13]. As an example in their paper they use a choice between transportation modes for the Northeast Corridor: the development of a high-speed rail transportation versus the VTOL (vertical take-off and landing) aircraft system. Their super-goal is "The Good Life" and as a means to that end they suggest four sub-goals: Convenience, Safety, Aesthetics, and Economics. Since these subgoals are rather vague, they are clarified by introduction of various specifications and sub-specifications. For example, Aesthetics is broken down into Users and Non-Users; and the User category is further broken down into Comfort, Noise, Visual. Figure 7 exhibits a slightly modified version of a figure in the Manheim-Hall paper.

As you notice in Figure 7, Manheim and Hall refer to "value-wise" dependent and independent categories. This dichotomy will play a central role in the sequel and gradually our definition of these concepts will be sharpened as our analytical demands for preciseness become more imperative. For the time being we quote from Manheim and Hall:

"Value-wise dependent goals are those that can be evaluated only in conjunction with other goals. An example of this could be in 'safety': if fatalities are very low, we may be willing to accept more injuries than we would if fatalities were higher. Value-wise independent goals, on the other hand, can be evaluated on their own without regard to any other goals."



KEY

--- SPECIFICATION

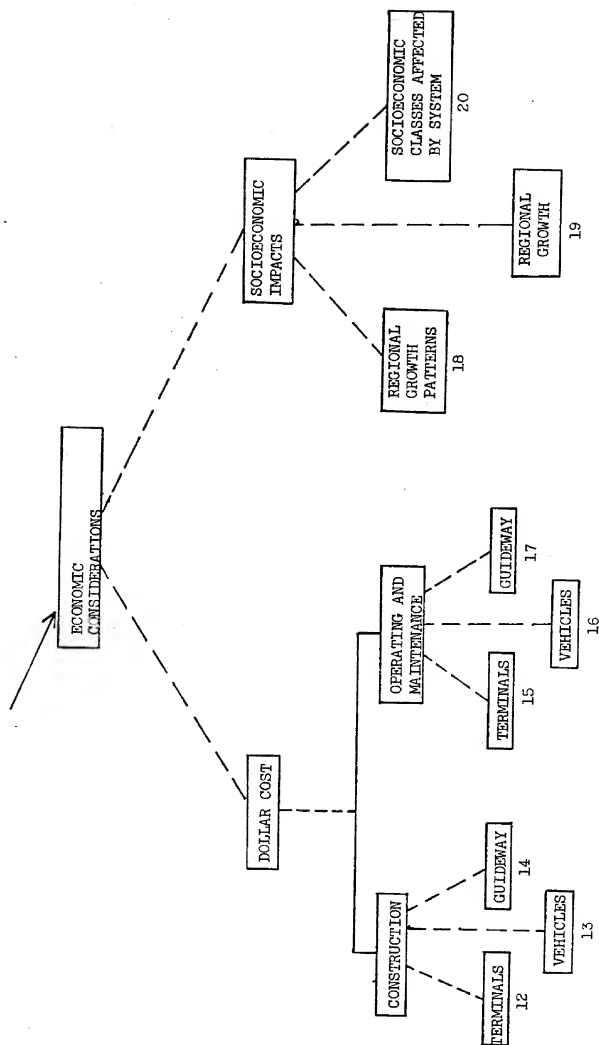
--- MEANS - END

< VALUE-WISE INDEPENDENT

[] VALUE-WISE DEPENDENT

Figure 7

RESULTS OF THE GOAL FABRIC ANALYSIS - PART I



KEY
 --- SPECIFICATION
 --- MEANS - END
 ^ VALUE-WISE INDEPENDENT
 ^ VALUE-WISE DEPENDENT

Figure 7

RESULTS OF THE GOAL FABRIC ANALYSIS - PART II

In explicating the super-goal of The Good Life for the transportation problem in the Northeast Corridor, the authors, by means of the specification and means-ends processes, characterize a plan in terms of twenty predictable goals or what I call "attributes". They are listed in Table 1.

TABLE 1

x_1	: Travel times (door to door)
x_2	: Probability of delay
x_3	: Out-of-pocket cost (door to door)
x_4	: Fatalities
x_5	: Non-fatal injuries
x_6	: Property damage
x_7	: User comfort
x_8	: User noise
x_9	: User visual
x_{10}	: Non-user visual
x_{11}	: Non-user noise
x_{12}	: Terminal construction cost
x_{13}	: Vehicle construction cost
x_{14}	: Guideway construction cost
x_{15}	: Terminal operating cost
x_{16}	: Vehicles operating cost
x_{17}	: Guideway operating cost
x_{18}	: Regional growth patterns
x_{19}	: Regional growth rate
x_{20}	: Socioeconomic classes affected

There is a real substantive question that I would like to duck even though it is of major importance; that is: Are these 20 attributes a sufficiently rich and meaningful set of descriptors to capture the essence of the problem area?

Now in making a paired comparison between one mode of rail transportation (R) and VTOL (V), the authors suggest that the decision maker should rank the alternatives R and V with regard to each of the 20 attributes, and then they display these preferences in a tree-like diagram; see Figure 8.

From here on the authors are a bit vague about where to go next. Of course, if R is better than V for all twenty attributes, then by dominance the choice is easy. If, however, on some attributes R is better and on others V is better, then there is a tug-of-war and somehow the decision maker must subjectively commensurate these objectively incommensurable attributes. The authors prefer not to push the formalism this far. They essentially end up saying that once the goal fabric is illustrated as in Figure 8, the decision maker must synthesize the information in his own mind and come up with a recommendation. In this paper I would like to start the analysis in the spirit of Manheim and Hall but I would prefer to push the formal apparatus much further than they do.

Let me now summarize this section by drawing together those suggestions made by Manheim and Hall that will fit in with the scheme I have in mind. In a given problem domain, the attributes describing a consequence are not given a priori to the analyst but they have to be created. I find the Manheim-Hall approach of looking at the hierarchy of goals, of specifications

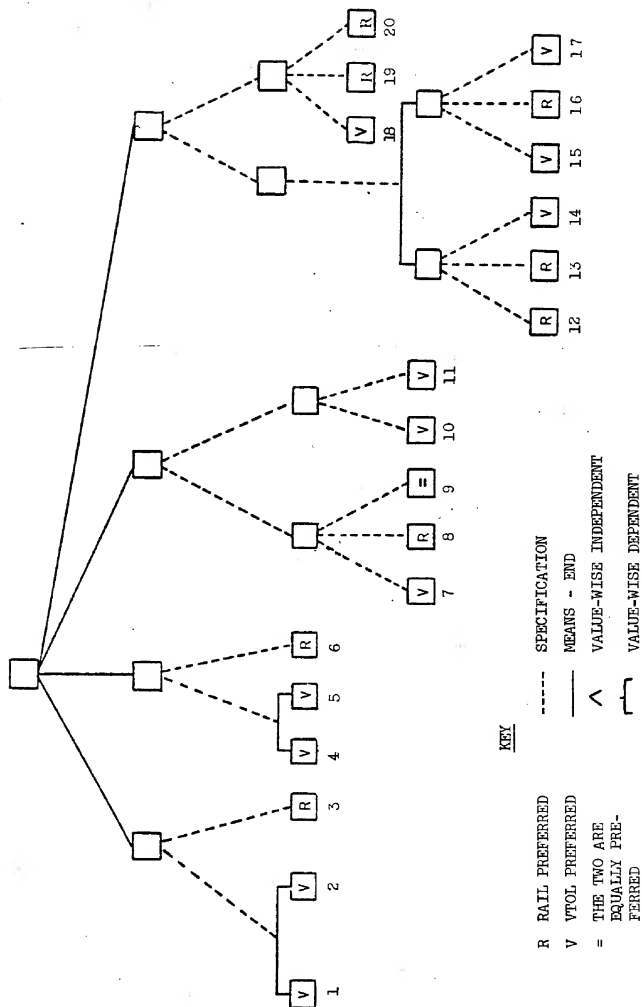


FIGURE 8

COAL FABRIC, SHOWING DOMINANCE AT PREDICTABLE LEVELS OF GOALS

of goals, and of means-ends relationships a desirable first step. Incidentally, the Miller paper [14] proceeds the same way initially. By means of this procedure the analyst constructs a set of attributes (x_1, x_2, \dots, x_r) , which can be thought of as descriptors that can effectively elaborate a given consequence or outcome. Thus for any consequence C the analyst can associate the r -tuple of evaluations $[x_1(C), \dots, x_r(C)]$, where $x_i(C)$ is a scaled value of C on the i th attribute. In some circumstances $x_i(C)$ may be a physical quantity like a travel time in minutes or an out-of-pocket cost in dollars; in other circumstances $x_i(C)$ may be a subjectively scaled value, like the subjective impression of comfort of a ride on a 5-point scale going from poor to excellent. It is not really critical at this point that $x_i(C)$ be given in numerical terms although for later manipulations it will be helpful if $x_i(C)$ gives a simple, clear evaluation of C on the i th attribute.

The rest of the paper will be concerned with what we can do with these r -tuples once we have them. However, insofar as the elaboration of a goal fabric is not unique, the analyst would be well-advised to choose a goal fabric to help facilitate subsequent analyses. But we cannot talk about this until we first discuss these possible subsequent evaluations.

5.3. Conditional Preferences

Let's simplify our notation a bit. Instead of writing $[x_1(C), x_2(C), \dots, x_r(C)]$ for the evaluation of a generic consequence C , we shall suppress the C and let the generic evaluation be $\underline{x} = [x_1, x_2, \dots, x_r]$, which we shall identify with a point in r -space. If we want to talk about several of these points, we shall use superscripts: $\underline{x}^{(1)}, \underline{x}^{(2)}, \dots$, or simply super primes $\underline{x}', \underline{x}'', \dots$. Our task is therefore to define -- or better yet to "create" -- a utility function, u , defined over points in r -space.

Let I designate the index set $\{1, 2, \dots, r\}$ and let Y be a subset of I . To keep a simple example in mind as we go along, let $r = 5$ and $Y = \{1, 2, 4\}$ so that the complement of Y , written \bar{Y} , would be $\bar{Y} = \{3, 5\}$. Now for any consequence C we might be concerned primarily with those attributes in Y and only secondarily with those attributes in \bar{Y} . Or alternatively we might want to decompose the set of attributes into those in Y and those in \bar{Y} according to a natural goal fabric. If \underline{x} is a generic point we shall decompose \underline{x} by considering those components of \underline{x} corresponding to attributes in Y and those in \bar{Y} . For notational convenience we shall designate \bar{Y} by Z . In order to avoid too many subscripts and superscripts we shall let \underline{y} represent components of \underline{x} corresponding to Y and \underline{z} components of \underline{x} corresponding to \bar{Y} or Z . Thus, in our illustrative example where $Y = \{1, 2, 4\}$ and $Z = \{3, 5\}$,

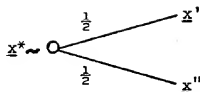
$$\underline{y} = \{x_1, x_2, x_4\} \quad \text{and} \quad \underline{z} = \{x_3, x_5\};$$

we shall also write: $\underline{x} = (\underline{y}, \underline{z})$.

Often we want to compare different \underline{y} 's when \underline{z} is held fixed at some value \underline{z}^0 say. If, for example,

$$\underline{x}' = (\underline{y}', \underline{z}^0) \quad \text{and} \quad \underline{x}'' = (\underline{y}'', \underline{z}^0)$$

and if \underline{x}' is preferred to \underline{x}'' (i.e., if $\underline{x}' > \underline{x}''$), then we will say \underline{y}' is conditionally preferred to \underline{y}'' given \underline{z}^0 . As another example suppose \underline{x}' , \underline{x}'' are as above, $\underline{x}^* = (\underline{y}^*, \underline{z}^0)$, and \underline{x}^* is indifferent to a lottery which yields \underline{x}' and \underline{x}'' with equal probability, i.e.,



In this case we shall say that x^* is conditionally indifferent to a 50-50 chance at x' and x'' given z^0 .

5.4. Utility Independence and Decomposition

Let us continue with the notation of the preceeding section. We shall be concerned in this section with the important case in which conditional preferences for y -components do not depend on the given z^0 . This can be formalized in terms of a weak and a strong form:*

Definition 1. Weak Conditional Utility Independence. The subset Y is weakly conditionally utility independent (wcui) of Z if and only if: (for all y', y'', z^0, z')

$$(y', z^0) \succeq (y'', z^0) \Rightarrow (y', z') \succeq (y'', z') ;$$

i.e., if conditional (ordinal) preferences on y -components do not depend on the given value of the common z -component.

Definition 2. Strong Conditional Utility Independence. The subset Y is strongly conditionally utility independent (scui) of Z if and

* We use the symbol $x' > x''$ to denote x' is strictly preferred to x'' , the symbol $x' \succeq x''$ to denote x' is preferred or indifferent to x'' , and the symbol $x' \sim x''$ to denote x' is indifferent to x'' .

only if the following holds: Let lotteries l' and l'' have consequences of the form (y, \underline{z}^0) where the second component is identical for all consequences. Then preferences for l' versus l'' do not change if the common value \underline{z}^0 is changed to \underline{z}' (for any \underline{z}^0 and \underline{z}').

We shall often use a loose interpretation of the above and say that the y component is wcui (or scui) the \underline{z} component and abbreviate this as: Y wcui Z (or Y scui Z).

If weak conditional utility independence holds, $(Y$ wcui $Z)$ we can legitimately talk about ordinal preferences for y 's without specifying any \underline{z} . And furthermore, we can attempt to work systematically on simplifying the y component. If Y wcui Z and if $y = (y_1, y_2)$, then we can legitimately talk about tradeoffs and substitution rates between components y_1 and y_2 without bringing \underline{z} 's into the picture. For example, suppose a professor has an offer of two academic jobs. He might consider the tradeoffs between teaching hours and quality of students without having to worry at the same time about the quality of nursery schools for his young daughter. [But even here, in some circumstances, weak independence may be destroyed if a shorter teaching load enables the professor to earn more money by consulting and thereby to afford a better nursery school for his daughter.]

The wcui and scui relationships are directional: we could have Y wcui Z but not Z wcui Y . For example, for y' , y'' , z' , z'' we might have

$$(y', z') > (y'', z') \quad \text{and} \quad (y', z'') > (y'', z''),$$

(i.e., $y' > y''$ for z' and z'') but

$$(y', z') > (y', z'') \quad \text{and} \quad (y'', z'') > (y'', z') ,$$

(i.e., $z' > z''$ for y' but $z'' > z'$ for y''). To be a bit more concrete suppose the DM (a salesman) is evaluating job offers and he analyzes each potential consequence in terms of a large number of attributes (x_1, x_2, \dots, x_r) . Let y denote the first component x_1 which gives his monetary reward (salary plus commissions) and let \underline{z} denote all other components. His ordinal preferences for different \underline{z} 's might very well depend on his financial level so that he can not assert that $Z \underline{w}cui Y$. But regardless of the determination of \underline{z} he would prefer more y to less y . In other words, $Y \underline{w}cui Z$. However, the salesman's attitude towards risky monetary gambles may depend on \underline{z} so that for him, although he agrees that $Y \underline{w}cui Z$, he may not assert that $Y \underline{s}cui Z$. Of course, the relation scui implies wcui but as the above example shows the converse is not true.

There is a special case that occurs often enough in practice that it warrants a few words. Let the first component x_1 of an r -tuple (x_1, x_2, \dots, x_r) designate a monetary payoff and assume that the following structure holds: for any $j \neq 1$ let $\underline{y} = (x_1, x_j)$ and \underline{z} denote all components $k \neq 1$ or j ; let $Y \underline{w}cui Z$. In this case we can examine the tradeoffs between x_1 and x_j without considering \underline{z} . Now we can "price-out" x_j by bringing x_j to zero (or to some other base position x_j^* , say) and by altering x_1 so as to maintain indifference. Under the assumptions we are working with, this procedure can now be iterated. First price out x_2 , then x_3 , then x_4 , and so on. In this fashion we can transform all x_k components, for $k \neq 1$, to a base case and be left with just the first component to consider when we convert r -tuples to utilities. This situation is a special, structured case of the general

procedure described in Section 4.3.

I mentioned before that the goal fabric, as discussed in Section 5.2, is not unique. The analyst should, wherever possible design his attribute space to exploit wcui and scui relationships.

5.5. Examples from Medical Treatment

In recent years several exploratory decision analyses of medical treatment problems have been done ranging from the treatment of the common sore throat to surgical treatments of duodenal ulcers.* None of these analyses (with the possible exception of the paper by Ginsberg and Offensend) was meant to be definitive. In each case there were branches of the decision tree that were not included because they would only have complicated matters and not have contributed to an understanding of the methodological issues involved in this approach. The probability assignments at chance nodes were only roughly assessed, and these numbers were not meant to be taken very seriously. In exploratory, pilot studies of this kind, however, these inadequacies are not serious stumbling blocks; more comprehensive trees could easily be drawn that would capture the essence of the problem, and a panel of experts could be convened to study published

* Dr. J. Polissar, after receiving his M.D. degree, continued working for a Ph.D. degree in Operations Research at Harvard. As part of a course he wrote a term paper on the "Treatment of the Common Sore Throat," Ronald Rubel, a doctoral student of mine at the Harvard Business School, considerably amplified Polissar's treatment and completed a thesis on "Decision Analysis in Medical Diagnosis and Treatment." Rubel worked on the problem of renal hypertension as well. Mr. Lewinnek, a senior medical student at Harvard, delivered a paper at the Boylston Society on the surgical treatment of duodenal ulcers and asymptomatic gall-stones. Dr. R. Greenes, another M.D. who is studying for his Ph.D., wrote a paper for me on "Diagnosis and Treatment of a Gastric Ulcer." Another paper in the same genre is by Ginsberg and Offensend, "An Application of Decision Theory to a Medical Diagnosis-Treatment Problem, p-3786, The RAND Corporation, 1968.

statistical results and assess responsible probabilities at chance moves. In each case the weakest link in the chain of the analysis was the treatment of the utility structure. I would now like to look at this problem more closely.

A typical consequence at the end of a branch of the tree in any one of these studies might be effectively summarized by a 7-tuple

(x_1, x_2, \dots, x_7) , where

x_1 = amount of money spent for treatment, drugs, etc.

x_2 = number of days in bed with a High index of uncomfortableness.

x_3 = number of days in bed with a Medium index of uncomfortableness.

x_4 = number of days in bed with a Low index of uncomfortableness.

$x_5 = \begin{cases} 1 & \text{if complication A occurs} \\ 0 & \text{does not occur} \end{cases}$

$x_6 = \begin{cases} 1 & \text{if complication B occurs} \\ 0 & \text{does not occur} \end{cases}$

$x_7 = \begin{cases} 1 & \text{if complication C occurs} \\ 0 & \text{does not occur} \end{cases}$

In the sore throat case complication A might be "death by shock" because of a penicillin injection reaction, B might be rheumatic fever, C might be glomerulonephritis (a very serious kidney disorder). The task is to assess a utility function over the 7-tuple $\underline{x} = (x_1, \dots, x_7)$.

It will simplify the presentation if we put all the awesome complications at the end of the tree. Thus let us group together all points with a common (x_1, x_2, x_3, x_4) history; these points are $(x_1, x_2, x_3, x_4, 0, 0, 0)$, $(x_1, x_2, x_3, x_4, 1, 0, 0)$, $(x_1, x_2, x_3, x_4, 0, 1, 0)$, and $(x_1, x_2, x_3, x_4, 0, 0, 1)$. We are assuming that double or triple complications are ruled out. These can be displayed on terms of the Chance move in Figure 9. At the node marked Z in this figure, we could summarize the entire Chance move by the generalized 7-tuple

$$(x_1, x_2, x_3, x_4, p_A, p_B, p_C)$$

where p_A , p_B , and p_C are respectively the conditional probabilities of the awesome complications A, B, and C at that particular branch of the tree.

It is reasonable to assume the following:

1) The trade-offs between x_1 , x_2 , x_3 , and x_4 (i.e., ordinal preferences amongst the first four components) do not depend on the particular values of p_A , p_B , and p_C .

2) The trade-offs between p_A , p_B , and p_C do not depend on the particular values of x_1 , x_2 , x_3 , and x_4 .

(In other words the index sets $\{1, 2, 3, 4\}$ and $\{5, 6, 7\}$ are mutually weakly conditionally utility independent.)

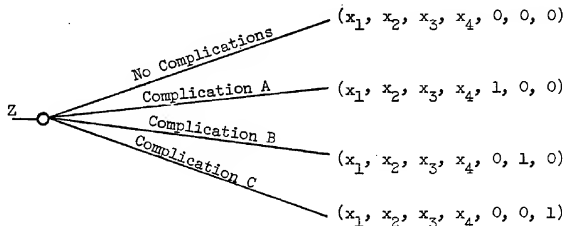


Figure 9

Now let us think hard about the 4-tuple (x_1, x_2, x_3, x_4) and try to summarize this 4-tuple in terms of a single numeraire -- let's say in terms of equivalent days in bed with medium discomfort (e.g., with a discomfort level described by 100° fever, headache, general weakness, slight nausea, little specific pain). We could go through the following three successive reductions:

$$(x_1, x_2, x_3, x_4) \rightarrow (x_1, 0, x_3', x_4) \rightarrow (x_1, 0, x_3'', 0) \rightarrow (0, 0, x_3''', 0).$$

For example, the first reduction asks: keeping x_1 and x_4 fixed, if x_2 is reduced to zero, how should x_3 be changed so that the modified 4-tuple is indifferent to the original 4-tuple? These successive reductions are not easy to do, but nevertheless it is possible to think about these reductions in a responsible, systematic, and meaningful way. Let us now summarize this special 4-tuple, with three zeroes, by means of a single numeraire, y say -- so that in the above special case y would be the number x_3''' .

Now let us think hard about the 3-tuple (p_A, p_B, p_C) and try to summarize this 3-tuple in terms of a single numeraire -- let's say in terms of equivalent probability of death (quick and without pain) -- ugh! Letting N represent no complications, let us assume that our preferences go:

$$A < B < C < N.$$

The stimulus (p_A, p_B, p_C) represents the Chance move in Figure 10. Now let us substitute for B and C equilibrating lotteries involving A and N as reference payoffs. Suppose that if the patient had complication B he

would just be willing to risk a π_B -chance at A to get a $(1 - \pi_B)$ -chance at N. Similarly for C replacing B. Making these substitutions we get that

$$(p_A, p_B, p_C) \sim (p_A + \pi_B p_B + \pi_C p_C, 0, 0) .$$

Let us now summarize the special 3-tuple, with two zeroes, by means of a single numeraire, z say -- so that in the above special case z would be the number $p_A + \pi_B p_B + \pi_C p_C$.

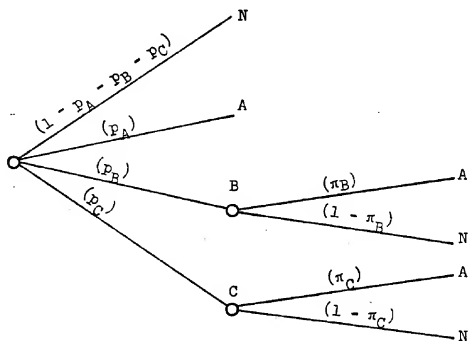


Figure 10

We have now reduced the 7-tuple (x_1, \dots, x_7) to the pair (y, z) where y is the pseudo-standardized days in bed and z is the pseudo-effective probability of complication A. Now comes the hardest part: How can one assess a utility function on (y, z) pairs?

In order to find substitution rates between y and z we are led to consider such questions as: "How many additional days in bed would you be willing to spend in order to reduce the probability of death from .001 (say) to 0?" Or, if the first component referred to monetary assets, the question might be posed as: "What proportion of your present assets are you willing to exchange in order to reduce the probability of death from .001 (say) to 0?"

I would like to defer analysis of these complicated questions to Section 8, since I shall need to develop some basic concepts first. But before turning to other matters perhaps you are wondering about who is supposed to be the decision maker in these medical problems. Is it the doctor or the patient? I would think that in most situations the decision maker is the doctor but ideally he should be responsive to the fundamental preferences of his patient. Now I certainly don't have in mind asking a miserable, feverish, apprehensive patient who has a severely inflamed sore throat whether he would rather be ill for an extra couple of days in order to reduce the chance of nephritis from .007 to .006. That would be silly indeed. However, it might not be unreasonable for medical researchers to investigate how a sample of individuals might feel about certain critical tradeoffs. Then, on the basis of these sampled responses and on a detailed decision analysis of a given medical problem,

these researchers might convey their qualitative findings to the general practitioner by means of articles, lectures, textbooks, and so forth.

6. Additive Representations

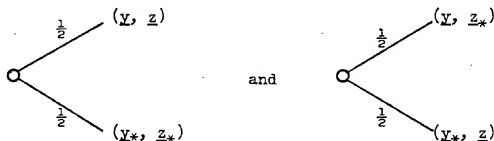
6.1. Acknowledgments

The material in this section draws heavily on the work of Miller [14] and Fishburn [2, 3, 4, 5, 6].

6.2 Case of Two Summands

Suppose, as we did previously, we let $\{Y, Z\}$ be a partition of the set of attributes, and let $\underline{x} = (y, z)$. Suppose $y_* \precsim y$ for any z and $z_* \precsim z$ for any y . Suppose that the DM wishes to assess a utility function u on \underline{x} -space, and that he is willing to adopt the following

Marginality Assumption: For any y, z , he feels indifferent between the following two lotteries:



Note that in each lottery he has a 50-50 chance between y and y_* and between z and z_* . Or stated somewhat differently, note that each lottery has the same marginal probability distribution for y -outcomes and for z -outcomes; the joint distributions differ, however.

By the assumed existence of a utility function u and the Marginality Assumption, we have

$$\frac{1}{2}u(y, z) + \frac{1}{2}u(y_*, z_*) = \frac{1}{2}u(y, z_*) + \frac{1}{2}u(y_*, z), \quad (1)$$

where we take the liberty to write $u(y, z)$ instead of $u((y, z))$.

There is now no loss of generality in letting

$$u(y_*, z_*) = 0, \quad (2a)$$

and defining

$$u_Y(y) = u(y, z_*), \quad (2b)$$

and

$$u_Z(z) = u(y_*, z). \quad (2c)$$

Substituting (2a, b, c) into (1), we get the additive representation

$$u(y, z) = u_Y(y) + u_Z(z). \quad (2d)$$

This result, although expositied differently here, is due to Fishburn.

Now let's see how we can use (2a, b, c) in obtaining measurements.

Clearly we must have $u_Y(y_*) = 0$ and $u_Z(z_*) = 0$, which sets the origins of measurement, but still we have a scale problem. In obtaining the u_Y function the decision maker can think of z_* as being held fixed and by asking questions about certainty equivalents for y -gambles he can determine u_Y up to a multiplicative-constant. The same goes for u_Z . But we cannot be sloppy about the relative scaling between Y and Z and this topic is of such crucial importance that it warrants a special subsection by itself. But first let us squeeze out some more results from the representation (2d).

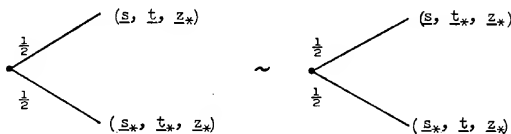
From (2d) we see that we have $Y \text{ scui } Z$ and $Z \text{ scui } Y$. Later on we shall show the converse does not hold, namely: $Y \text{ scui } Z, Z \text{ scui } Y$, and the existence of a u-function do not collectively imply the additive representations in (2d). But still these properties can and will be exploited later on.

6.3. Generalization to More Summands

Let $\{Y, Z\}$ be a partition of I -- i.e., $Y \cup Z = I$ and $Y \cap Z = \emptyset$. Assume, as in the previous section, that u has a representation in the form

$$u(\underline{y}, \underline{z}) = u_Y(\underline{y}) + u_Z(\underline{z}).$$

Now partition Y into the subsets S and T -- i.e., $S \cup T = Y$, $S \cap T = \emptyset$ -- and represent \underline{y} in terms of $(\underline{s}, \underline{t})$ in the natural way. Let $\underline{y}_* = (\underline{s}_*, \underline{t}_*)$. Now if the DM is indifferent between the following two lotteries



-- again we will refer to this as a "marginality assumption" -- then using the same proof as we did in the last subsection we get

$$u_Y(\underline{s}, \underline{t}) = u_S(\underline{s}) + u_T(\underline{t}), \quad (3a)$$

where

$$u_S(\underline{s}) \equiv u_Y(\underline{s}, \underline{t}_*) \quad (3b)$$

and

$$u_T(t) = u_Y(s_*, t) . \quad (3c)$$

Combining (2d) and (3a) we get

$$u(s, t, z) = u_S(s) + u_T(t) + u_Z(z) , \quad (4)$$

and now the generalization should be clear. In terms of the goal fabric discussed in Section 5.2 we could think of Y and Z as a partition of our super-goal, and S and T as a partition of the goal Y. Obviously this procedure could be extended throughout a hierarchical arrangement -- provided, or course, marginality assumptions are appropriate.

Let us now summarize by stating a more symmetrical form of the main theorem due to Fishburn. Let $\{S_1, S_2, \dots, S_k\}$ be a partition of I -- i.e., $S_1 \cup S_2 \cup \dots \cup S_k = I$, and $S_i \cap S_j = \emptyset$ for all $i \neq j$. Let $\underline{x} = (\underline{s}_1, \underline{s}_2, \dots, \underline{s}_k)$ where \underline{s}_i represents the components of \underline{x} corresponding to S_i . Let Y be the union of any number of the S_i 's; let $Z = \bar{Y}$; and let $\underline{x} = (\underline{y}, \underline{z})$ in the natural way. If for any $\{Y, Z\}$ the marginality assumption holds, then we have

$$u(\underline{s}_1, \dots, \underline{s}_k) = u_1(\underline{s}_1) + u_2(\underline{s}_2) + \dots + u_k(\underline{s}_k) \quad (5a)$$

where

$$u(\underline{s}_{1*}, \underline{s}_{2*}, \dots, \underline{s}_{k*}) = 0 \quad (5b)$$

and

$$u_i(\underline{s}_i) = u(\underline{s}_{i*}, \dots, \underline{s}_{i-1*}, \underline{s}_i, \underline{s}_{i+1*}, \dots, \underline{s}_{k*}) \quad (5c)$$

for $i = 1, 2, \dots, k$. Notice that for typographical convenience I use u_i instead of the more cumbersome notation u_{s_i} .

6.4. Relative Scaling

Let's continue with the notation of the previous subsection. Assume that marginality assumptions hold and that u has the representation given in (5). For each i choose a value s_i^* such that any s_i is in between s_{i*} and s_i^* (i.e., $s_{i*} \preceq s_i \preceq s_i^*$). This involves some loss of generality but in practice this is not a very binding assumption. In the form (5a) we do not have the freedom (because of scaling problems) to set $u_i(s_i^*)$ equal to any specified value but let us set $u_i(s_i^*) = 1$ nevertheless and compensate for this choice by inserting a scaling or weighting factor λ_i . In other words, let us use the representation:

$$u(s_1, s_2, \dots, s_k) = \lambda_1 u_1(s_1) + \lambda_2 u_2(s_2) + \dots + \lambda_k u_k(s_k) \quad (6a)$$

where

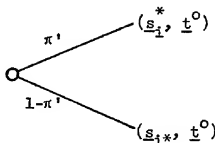
$$u(s_{1*}, s_{2*}, \dots, s_{k*}) = 0, \quad u(s_1^*, s_2^*, \dots, s_k^*) = 1, \quad (6b)$$

$$u_i(s_{i*}) = 0, \quad \text{and} \quad u_i(s_i^*) = 1 \quad (6c)$$

and where $u_i(s_i)$ is the conditional utility function on the s_i -component, scaled in accordance with (6c). We can talk about the u_i function because of the scui property.

Before going on with a discussion of the λ_i 's let's take a closer look at the measurement problem involved in the determination of u_i . Imagine that all the s_j 's for $j \neq i$ are held fixed at some specific value; call this common value by \underline{t}^0 and consider \underline{x} 's of the form (s_i, \underline{t}^0) . We

now normalize u_i such that $u_i(\underline{s}_{i*}, \underline{t}^0) = 0$, $u_i(\underline{s}_i^*, \underline{t}^0) = 1$. If a particular value \underline{s}_i^* and a number π' are such that the SM is indifferent between getting $(\underline{s}_i^*, \underline{t}^0)$ for certain and getting the lottery



(i.e., the lottery which yields \underline{s}_i^* with probability π' and \underline{s}_{i*} with probability $1-\pi'$ -- it always being understood that the same \underline{t}^0 represents all components other than those in \underline{s}_i), then, by definition,

$$u_i(\underline{s}_i^*) = \pi'_i.$$

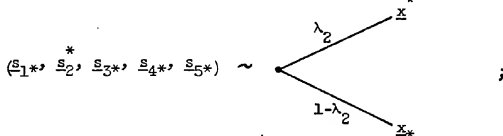
Now that we have adopted a convention for the scales of the u_i 's we can turn our attention to the λ_i 's. A Basic Reference Lottery Ticket which yields $\underline{x}^* = (\underline{s}_1^*, \dots, \underline{s}_k^*)$ with probability λ and $\underline{x}_* = (\underline{s}_{1*}, \dots, \underline{s}_{k*})$ with probability $1-\lambda$ will be referred to as λ -brlt. Thus a 1-brlt is merely \underline{x}^* , a 0-brlt is \underline{x}_* , and a 0.5-brlt gives an equal chance at \underline{x}^* and \underline{x}_* . Now for any subset T of the indices $\{1, 2, \dots, k\}$, let $X(T)$ be the \underline{x} point where the i th factor is \underline{s}_i^* if i belongs to T and is \underline{s}_{i*} if i does not belong to T . Thus if $k = 5$ and $T = \{1, 2, 4\}$ then

$$X(T) = X(1, 2, 4) = (\underline{s}_1^*, \underline{s}_2^*, \underline{s}_3^*, \underline{s}_4^*, \underline{s}_5^*) .$$

In terms of this symbolism $X(\phi) = \underline{x}_*$ and $X(I) = \underline{x}^*$, where I is the index set $\{1, \dots, k\}$. One more definition: Let

$$\lambda(T) = \sum_{i \in T} \lambda_i .$$

Now prepared with this notational armory we can proceed. First, λ_1 is such that $X\{i\}$ is indifferent to a λ_1 -brlt, for $i = 1, 2, \dots, k$. (To illustrate, let $k = 5$; then λ_2 , say, will be such that



i.e., $X\{2\}$ is indifferent to a λ_2 -brlt.)

These k relations define the λ_i 's and can be used for assessment purposes. But one must keep in mind the consistency requirements:

- a. $\lambda_i \geq 0$, all i .
- b. $\sum_{i=1}^k \lambda_i = 1$.

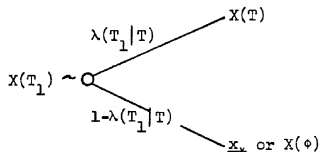
From the individual λ_i 's we can now obtain $\lambda(T)$ for each subset T . However, just as in the measurement of discrete probabilities it may not necessarily be easiest to proceed in this manner. It may be preferable for the decision maker to think about assigning a value directly to $\lambda(T)$ for subsets T containing more than one element and thereby to impose a

condition that the λ_i 's are such that

$$\sum_{i \in T} \lambda_i = \lambda(T) .$$

Notice that $\lambda(T)$ is the equilibrating value such that $X(T)$ is indifferent to a $\lambda(T)$ -brlt.

In assigning probabilities to a finite set of mutually exclusive and collectively exhaustive elementary events $\{E_1, E_2, \dots, E_T\}$ it is often natural to make an assignment first to a subset of elementary events and then to use conditional probability considerations to further subdivide this assignment. We might find it helpful to proceed in an analogous manner in the present context. To this end suppose T_1 is a subset of T . We could then ask what portion of the weight $\lambda(T)$ should be assigned to T_1 . Letting $\lambda(T_1|T)$ be the equilibrating value such that



and substituting a $\lambda(T)$ -brlt for $X(T)$, we establish the rule that

$$\lambda(T_1) = \lambda(T_1|T) \cdot \lambda(T), \quad \text{for } (T_1 \subset T) ,$$

which is analogous to the multiplication rule of probability theory.

How one finally chooses to assess the λ_i 's, whether directly or indirectly by first assigning λ -values to a union of S_i 's and then conditional λ -values, depends on which procedure seems most natural in the context of the real problem under consideration.

One final word of caution: it is not easy to interpret the λ_i 's since they depend on the choices $(\underline{s}_{1*}, \underline{s}_1^*; \dots; \underline{s}_{k*}, \underline{s}_k^*)$. Thus if $\lambda_1 = .1$ and $\lambda_2 = .2$ we cannot say that attributes S_2 are twice as important as attributes S_1 ; indeed we cannot even say that attributes S_2 are more important than S_1 . We can say, however, that if, starting from the point $(\underline{s}_{1*}, \underline{s}_{2*}, \dots, \underline{s}_{k*})$, we would rather raise \underline{s}_{2*} to \underline{s}_2^* than raise \underline{s}_{1*} to \underline{s}_1^* , then $\lambda_2 > \lambda_1$. More generally, if we would rather raise all the \underline{s}_{i*} 's to \underline{s}_i^* 's for all i belonging to subset T_2 than for subset T_1 , then $\lambda(T_2) > \lambda(T_1)$. If in a loose sense \underline{s}_{i*} and \underline{s}_i^* are "close together" -- i.e., if there is not much of a difference in the bounds on the \underline{s}_i component -- then the λ_i may be small but still the S_i attribute may be mighty important. Suppose, for example, in comparing jobs S_i refers to monetary rewards and all jobs under consideration pay almost the same amount so that \underline{s}_{i*} and \underline{s}_i^* are close together; then λ_i may be small but this does not mean that money is unimportant to the DM.

6.5. An Example

Let's use the example described in Section 5.2 (due to Manheim and Hall) and whose qualitative goal fabric is displayed in Figure 7. There are twenty attributes that play a role in that example but as the authors point out not all of these are what they call value-wise independent. My aim is now to partition these twenty attributes into S_i subsets that

are value-wise independent in the Manheim-Hall sense and I shall assume for the purpose of this illustrative example that the S_i subsets obey the marginal assumption needed for the additive representation given in (6). To this end, we define the S_i 's as in Table 2.

Table 2

S_1	= {1, 2}
S_2	= {3}
S_3	= {4, 5}
S_4	= {6}
S_5	= {7}
S_6	= {8}
S_7	= {9}
S_8	= {10}
S_9	= {11}
S_{10}	= {12, 13, 14, 15, 16, 17}
S_{11}	= {18}
S_{12}	= {19}
S_{13}	= {20}

Suppose, for the sake of concreteness, that there are only a dozen \underline{x} 's (consequences) to be evaluated: $\underline{x}^{(1)}, \underline{x}^{(2)}, \dots, \underline{x}^{(12)}$. We now score the j th consequence ($j = 1, 2, \dots, 12$) on the i th attribute ($i = 1, 2, \dots, 20$) and denote this value $x_i^{(j)}$. For some i this score will be an objective physical or engineering number, for other i

it may be a value on some psychological scale. These entries can be arranged as shown in Table 3. I am not sure what Mannheim and Hall have in mind for the 20th attribute, "socioeconomic classes affected," but let us duck this problem by assuming somehow this attribute has been elaborated by introducing various specifications and by some means or other the consequences have been given an overall score on attribute # 20. For each attribute i a range of scores is given on the right hand side of the table.

Table 3

Attributes	Consequences					Range	
	1	...	j	...	12	Worst	Best
1. Travel time (door to door) in hours	3.5	...	2.7	...	3.2	3.7	2.5
⋮	⋮		⋮		⋮	⋮	⋮
7. User comfort (10 point scale)	6	...	3	...	8	1	8
⋮	⋮		⋮		⋮	⋮	⋮
20. Socioeconomic classes affected (10 point scale)	4	...	9	...	0	0	10

We now can think of the column under "Worst" as the representation of \underline{x}_* and the column under "Best" as the representation of \underline{x}^* .

Referring back to the goal Fabric in Figure 7 we now can combine some of the S_i 's into meaningful higher order categories. This is done in Table 4. We now can meaningfully ask qualitative questions about $\lambda(T_i)$'s.

Table 4

	<u>Identification</u>
$T_1 = \{s_1, s_2\}$	Convenience
$T_2 = \{s_3, s_4\}$	Safety
$T_3 = \{s_5, s_6, s_7\}$	Aesthetics: User
$T_4 = \{s_8, s_9\}$	Aesthetics: Non-User
$T_5 = \{s_{10}\}$	Dollar Cost for Construction, Operating, and Maintenance
$T_6 = \{s_{11}, s_{12}, s_{13}\}$	Socioeconomic Impacts

For example: "Imagine that all evaluations are at the worst state. Would you rather have the attributes in T_1 pushed from worst to best than the attributes in T_2 pushed from worst to best?" [If yes, this means $\lambda(T_1) > \lambda(T_2)$.]

"Would you rather have the attributes in $T_1 \cup T_2 \cup T_5$ pushed from worst to best than the attributes in $T_3 \cup T_4 \cup T_6$ pushed from worst to best?" [If yes, this means $\lambda(T_1 \cup T_2 \cup T_5) > \lambda(T_3 \cup T_4 \cup T_6)$ which means $\lambda(T_1 \cup T_2 \cup T_5) > .5$.]

Hence without asking any probabilistic questions it is possible to get some qualitative "feel" for the $\lambda(T_i)$ -values. If there exists a subset T of attributes such that $X(T) \sim X(\bar{T})$, then we can infer that $\lambda(T) = \lambda(\bar{T}) = \frac{1}{2}$. And, more generally, if $X(T_1) \sim X(T_2)$, then $\lambda(T_1) = \lambda(T_2)$. In many ways the richer the set of attributes the easier it becomes to group attributes in a way that permits the analyst to infer properties of the λ -function without asking probabilistic questions. I take it that at any stage of an analysis such as this, the sophisticated analyst would

use sensitivity checks to determine whether he need probe any further. Perhaps the crude qualitative measures already obtained suffice to resolve the problem.

Another methodological point that needs clarification is the notion of consistency: When questions are asked one way it might turn out that $\lambda(T_1) > .25$ say, and when asked another way $\lambda(T_1) < .25$. This will happen and when it does the DM will just have to think harder about the issues and modify some of his evaluations in order to obtain consistency. This is psychologically painful and time-consuming and once again this step should be preceded by a sensitivity analysis to determine whether the indeterminacy is worth resolving.

Let's go on. Suppose that we ascertain that $\lambda(T_1 \cup T_2 \cup T_5) = .6$, i.e., the DM is indifferent between $X(T_1 \cup T_2 \cup T_5)$ and a lottery that gives a .6 chance at the best prize x^* and a .4 chance at the worst prize x_* . Now we can ask conditional qualitative questions about $\lambda(T_1)$, $\lambda(T_2)$, $\lambda(T_5)$, $\lambda(T_1 \cup T_2)$, $\lambda(T_1 \cup T_5)$, $\lambda(T_2 \cup T_5)$. We could also ask probabilistic questions as: "Find a π value such that you are indifferent between $X(T_1)$ and between a lottery that gives you a π chance at $X(T_1 \cup T_2 \cup T_5)$ and a $(1 - \pi)$ chance at x_* (which is $X(\emptyset)$). Suppose we get $\pi = .3$. Then we have

$$\lambda(T_1 | T_1 \cup T_2 \cup T_5) = .3$$

and from these results we get

$$\lambda(T_1) = \lambda(T_1 | T_1 \cup T_2 \cup T_5) \cdot \lambda(T_1 \cup T_2 \cup T_5) = .3 \times .6 = .18.$$

Once we have $\lambda(T_1)$, and recalling that $S_1 \cup S_2 = T_1$ we can obtain $\lambda(S_1)$ and $\lambda(S_2)$ by asking questions that would yield $\lambda(S_1|T_1)$ and $\lambda(S_2|T_1)$.

Let's look closer at the u-evaluation of s_1 . Since $S_1 = \{1, 2\}$ and $s_1 = (x_1, x_2)$ we are now concerned with the utility evaluation over points in a two-space. Remember, we are not willing to accept a marginality assumption between these first two attributes so a simple additive scheme is not appropriate. We could, however, use the scheme outlined in Section 4.2 on variable substitution, which essentially reduces a two-dimensional representation to a unidimensional representation before the conversion to utilities is effected. Other schemes will be discussed in Section 7.

If we add the flesh and body tissue to this skeleton outline you can (I hope!) see how the DM would end up with a utility evaluation $u(\underline{x}^{(j)})$ for any consequence $\underline{x}^{(j)}$. Keep in mind that these u-values are now appropriate for probabilistic combinations. Thus, for example, if strategy, σ , will result in $\underline{x}^{(1)}$, $\underline{x}^{(2)}$ or $\underline{x}^{(3)}$ with probabilities .5, .3, and .2 respectively, then σ has an expected utility of

$$.5 u(\underline{x}^{(1)}) + .3 u(\underline{x}^{(2)}) + .2 u(\underline{x}^{(3)})$$

and this number is all that is necessary to know about σ when it is compared with other strategies -- of course, under the proviso that administrative and political considerations have been included in the \underline{x} -representations or are to be ignored altogether for the immediate purposes of the present analysis.

[Digression: It might be helpful at this point to mention briefly another approach to the multiattribute problem. Suppose that to each strategy choice, σ , of the DM there is an associated payoff (x_σ, y_σ) where preferences are for more x and more y . Since it is not possible, in general, to choose σ to maximize x_σ and y_σ simultaneously, one common approach is to set up some aspiration level x^* and try to maximize y_σ subject to the constraint $x_\sigma \geq x^*$. It is then appropriate to investigate the behavior of the maximum value as a function of x^* . In the case that x_σ and y_σ are random variables -- let's use super-tilde signs to remind us of this -- then it is meaningless to maximize \tilde{y}_σ subject to the constraint $\tilde{x}_\sigma \geq x^*$. One common procedure is to choose some high probability level like $\gamma = .95$ and to maximize the expected value of \tilde{y}_σ subject to the condition that

$$P(\tilde{x}_\sigma \geq x^*) \geq \gamma.$$

There are several obvious questions:

- a. Why use expected value of \tilde{y}_σ ?
- b. How can one think systematically about x^* and γ ?

In order to answer these questions one could be a bit more sophisticated and worry about the entire distribution of \tilde{y}_σ -- perhaps introducing a utility function -- and place some fancy constraint on the distribution of \tilde{x}_σ . But observe that so far we have only considered the marginal distributions of \tilde{x}_σ and \tilde{y}_σ . We have shown in this section, however, that if the DM is only concerned about these marginal distributions, then his utility function for

(x, y) pairs has an additive representation,

$$u(x, y) = u_X(x) + u_Y(y) ,$$

which makes the remaining measurement problem reasonably tractable.

6.6 Miller's Additive Representation of Worth

The techniques discussed in Sections 6.2 to 6.5 are in spirit, but not in detail, similar to procedures posed by Miller [14]. The main difference is that Miller's "worth" or "gratile" values are not meant to be utilities in the probabilistic sense. He associates a "worth" value to a vector of attribute scores in the form

$$W(x_1, \dots, x_r) = \sum_{i=1}^r \lambda_i W_i(x_i) ,$$

where the weights λ_i are constrained to be positive and to sum to unity and where the W_i -functions are constrained to lie in the interval from 0 to 1. From my point of view these λ_i weights and W_i functions in his system lack a clear-cut interpretation and the assessments appear a bit ad hoc. By embedding his system in a probabilistic choice context, we have given a standard utility interpretation to the quantities in his system and also we have given sufficient conditions to justify the additive representation he used.

It will be instructive to repeat here in skeleton form an example taken from Miller [14]. The context of the problem is the selection of a job by a recent college graduate; the subject is male and unmarried. This subject partitioned his considerations into four major goals: monetary compensation, geographical location, travel requirements, nature of work.

Each of these in turn were subdivided and some of these subdivisions were again subdivided. The goal fabric is exhibited in Figure 11. Each consequence is represented in terms of 15 attributes. The λ -weight for monetary compensation is .33 and this weight is broken down into a .70 to .30 split between immediate and future and these in turn are split in a manner shown in the figure. The weight assigned to attribute 1, for example, is

$$.33 \times .70 \times .90 = .209 .$$

There is a gnawing question whether these 15 attributes are indeed "value-independent." For example, the subject's attitude towards many aspects of geographical location might depend on his monetary compensation. Also one might wonder whether sex should get into the act. I am serious. One subject that I worked with in a similar analysis was primarily concerned with his chances of finding a suitable wife. I mention this just to point out a real concern: not only do we have to worry about value-independence of attributes but completeness of the set of attributes. Do they cover all the important issues? Miller talks about these questions and it would be inappropriate for me to paraphrase his extensive remarks.

Miller suggests that the DM should decompose each goal (a facet of his concern) into subgoals and these subgoals into "subsubgoals" until a level of detail is reached where objective, physical measurements can be associated with each attribute. For example, geographical location is broken down in his illustration to: proximity to relatives, degree of urbanity, and climate. Now when the DM thinks about "proximity to relatives", for

Figure 11 TOTAL WORTH	(13) Monetary Compensation	(17) Geographical Location	(17) Travel Requirements	(23) Nature of Work	(.20)	Mount. Training	.040	(15)	Direct Worth Estimate
					(.60)	(.30)	.059	(14)	Direct Worth Estimate
					(.50)	Technical Content	.049	(13)	Direct Worth Estimate
					(.40)	Immediate	.132	(12)	# Mos. Training
					(.60)	Trip Length	.082	(11)	Max. Trip Length
					(.80)	Long Trips	.054	(10)	% Time away/yr.
					(.40)	Proportion Time Away	.034	(9)	# Hrs/yr. 1-way travel
					(.20)	Daily Commuting	.034	(8)	Direct Worth Estimate
					(.20)	Climate	.004	(7)	SMA Population
					(.40)	Degree of Urbanity	.008	(6)	1-way Jet Flight time
					(.40)	Proximity to Relatives	.008	(5)	Loc. Adj. A.T. Dollars
					(.35)	10-Yr. Increase	.035	(4)	Loc. Adj. A.T. Dollars
					(.65)	3-Yr. Increase	.064	(3)	Loc. Adj. A.T. Dollars
					(.40)	Retirement	.009	(2)	Loc. Adj. A.T. Dollars
					(.10)	Fringe	.014	(1)	Loc. Adj. A.T. Dollars
(.80)	Insurance	.008	(1)	Loc. Adj. A.T. Dollars					
(.90)	Start Salary	.008	(1)	Loc. Adj. A.T. Dollars					

example, he might keep in mind the one-way jet time of the flight. But why stop there? He could also keep in mind comfort, cost of travel, convenience of schedules and so on. How far should this decomposition go? In all questions of this kind there is no simple answer. It depends on the personalities involved, the character and importance of the problem, the need of the DM to defend his actions, the availability of expert advice in some given facet of the problem, and so on. Also one can elect to decompose the problem down to a very fine partition of detail in qualitative terms but introduce quantitative evaluations only at a much broader level of aggregation. As an extreme case of this, one might choose to elaborate the detailed goal fabric of a problem, but refuse to assign utility numbers at any level of disaggregation below the super-goal of "the good life."

6.7 Use of a Classical, Additive-Utility Theorem

Economists of old, say way back in the 19th century, often based theoretical arguments on utility functions which were additive over the components of commodity bundles. For example, if commodity bundle #1 contained x_1^1 units of commodity 1 and bundle #2 contained x_1^2 units of commodity 1, for $i = 1, \dots, r$, and if a consumer chose bundle #1 rather than #2 then it would be asserted, "Bundle #1 is preferred to bundle #2 by the consumer because it has a higher utility index for him." They talked about "utility" as being a latent, underlying, psychological concept -- a sort of tickling of the synapses of the brain cells -- and it was because bundle #1 had a higher utility than #2 that it was chosen. Utility was treated as a cardinal concept and going from a utility of 5 to 10 was said to be equivalent to going from a utility of 10 to 15. Much of this traditional economic theory faded in importance when Pareto established that ordinal

(and not cardinal) utilities were adequate for most of the then existing economic theory. Be that as it may, I am interested now in examining a special case of classical (non-probabilistic) utility theory and seeing if it can be exploited for our present purposes.

A utility function V , defined on r -tuples, is a bona fide representation of the true tastes of the decision maker if r -tuple (x_1', \dots, x_r') is preferred or indifferent to (x_1'', \dots, x_r'') when and only when

$$V(x_1', \dots, x_r') \geq V(x_1'', \dots, x_r'') .$$

In this classical sense, there is no requirement that the expected utility number be an appropriate index to maximize in probabilistic choice situations.

In order not to lose sight of the point that the index V is not a probabilistic utility index (i.e., the expected value of V is not a guide to action), I chose to use the symbol V rather than U and henceforth I will refer to V as a "value" function.

A value function V defined on \underline{x} will be said to have an additive representation if there are real-valued functions v_1, \dots, v_r each defined on a single variable such that

$$V(x_1, \dots, x_r) = \sum_{i=1}^r v_i(x_i) .$$

In this section we shall discuss necessary and sufficient conditions for a value function to be additive and how such a function V can be used in analyzing choice problems with probabilistic outcomes.

Consider any r -tuple (x_1, x_2, \dots, x_r) ; suppose we hold fixed all components other than i and j , and ask the question: If we change x_i to x'_i , how much must we change x_j by so that the modified r -tuple is indifferent to the original r -tuple? This question involves substitution rates between components i and j , holding all other components fixed. In general, the substitution rates between i and j will depend on some of the components x_k for $k \neq i$ or j . But notice that when V is additive, the substitution rate between i and j does not depend on any of the values of components other than i and j . This, of course, is because in the V -representation the summands $v_k(x_k)$ for $k \neq i$ or j are all held fixed. Debreu [1] proved the profound and important result that the following converse is true, namely:

If for all i and j the substitution rate between the i^{th} and j^{th} components does not depend on the values of components other than i and j , if $r \geq 3$, (and if some innocuous continuity assumptions, which I don't wish to spell out here, hold), then there exist functions v_1, \dots, v_r and a bona fide value representation of the form

$$V(x_1, \dots, x_r) = \sum_{i=1}^r v_i(x_i) .$$

Debreu gave a non-constructive, topological proof of the above theorem. More recently Luce and Tukey [10], in an article on conjoint measurement, gave a more constructive, algebraic proof of Debreu's proposition.

Suppose now we know that the above representation theorem is true and that for any (x_1, \dots, x_r) it is appropriate to use the index

$\sum_{i=1}^r v_i(x_i)$ for ranking purposes. For our purposes, we now wish to investigate three questions:

1. How can one constructively assess the v_i -functions?
2. How can one use the V-representation in probabilistic choice situations?
3. What is the relation, if any, between the Debreu-independence assumption (i.e., the substitution rates between x_i and x_j do not depend on values of x_k for $k \neq i$ and j) and the Fishburn-marginality assumption (i.e., in probabilistic-choice situations only the marginal distributions of the attribute scores are of relevance)?

We will answer these questions in reverse order.

Suppose first that the Fishburn-marginality assumption is true and that a probabilistic utility function u exists. Then, as we have seen in Section 6.3, the u -function has an additive representation, viz:

$$u(x_1, \dots, x_r) = \sum_{i=1}^r u_i(x_i) .$$

Now if u is appropriate for probabilistic choice then, a fortiori, it is appropriate for non-probabilistic ranking of alternatives, and letting $v_i = u_i$ and $V = u$ we see that Fishburn-marginality implies Debreu-independence.

Next suppose that we start with Debreu-independence plus the existence of a probabilistic utility function u . Does this imply Fishburn-marginality? Can we get an additive representation for u ? The answer is No. To see this, suppose we let g be any monotone increasing function of a real variable and arbitrarily define

$$\begin{aligned} u(x_1, \dots, x_r) &= g[V(x_1, \dots, x_r)] \\ &= g[\sum_{i=1}^r v_i(x_i)] . \end{aligned}$$

Now if we let u govern probabilistic choices, then none of our assumptions are violated and clearly in this circumstance we will not, in general, be able to express u in additive form. This demonstration indicates that Debreu-independence is a necessary but not a sufficient condition for Fishburn-marginality. This leaves open the question of finding other easily verifiable and interpretable conditions which would co-imply Fishburn-marginality. This will be discussed in Section 7.

A very simple example might help at this point. Suppose the DM has three bank accounts and has an amount of money x_i in the i^{th} account. Would he rather have (x_1', x_2', x_3') or (x_1'', x_2'', x_3'') ? If all he cares about is the sum, then he can choose

$$V(x_1, x_2, x_3) = x_1 + x_2 + x_3 ,$$

and Debreu-independence obviously holds. But in risky-choice situations he can not in general find a utility function, u , such that

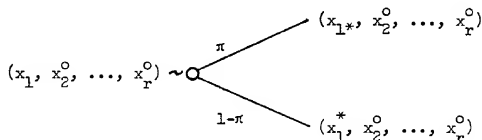
$$u(x_1, x_2, x_3) = u_1(x_1) + u_2(x_2) + u_3(x_3) .$$

This example, besides showing that Debreu-independence does not imply the Fishburn-marginality assumption, also illustrates a point that is often important in practice: when various x_i 's are in the same physical units often they should be first combined algebraically, and it may only needlessly complicate the analysis if one tries to convert each x_i separately into a value, worth, or "gratite" score.

Let us now turn to the second question posed above. Suppose that Debreu-independence is satisfied and V has an additive representation but u does not have additive representation. Can we use the V representation to construct a probabilistic utility function u ? A simple-minded approach, which will turn out to be not really operational, tries to assign a utility index u to each single summary measure V in a way that would reflect the decision maker's preferences in probabilistic choice situations. This would require answering such questions as: What value V for certain would you exchange for a 50-50 gamble at V_1 and V_2 ? The trouble with this approach is that the values of V are often not easily interpretable -- although this is not the case in the simple example cited in the preceding paragraph -- and it is difficult to reflect responsibly on gambles with abstract constructs as the prizes. We might, however, proceed as follows: Let

$$V(x_1^*, x_2^0, \dots, x_r^0) = V_a \quad \text{and} \quad V(x_1^*, x_2^0, \dots, x_r^0) = V_b.$$

By associating x_1 -values and π -values such that



we can determine the shape of the u -function for V values between V_a and V_b . In particular if x_1 and π are as depicted above and if $V(x_1^*, x_2^0, \dots, x_r^0) = V_c$ (say) then

$$u(V_c) = \pi u(V_a) + (1 - \pi)u(V_b) .$$

Thus if we commit ourselves to values for $u(V_a)$ and $u(V_b)$ we can ask reasonable questions to get values of $u(V)$ for $V_a \leq V \leq V_b$. Everything would be lovely now if the interval from V_a to V_b were sufficiently broad to cover the V -range we might be interested in. But suppose it is not broad enough. We then must seek ways to extend this range. We can proceed by holding $r-1$ components fixed at some other values and letting the remaining component vary. For example, suppose we restrict our attention to r -tuples of the form $(x_1^\dagger, x_2^\dagger, x_3^\dagger, \dots, x_r^\dagger)$ and suppose

$$V(x_1^\dagger, x_{2*}^\dagger, x_3^\dagger, \dots, x_r^\dagger) = V_d, \quad V(x_1^\dagger, x_2^*, x_3^\dagger, \dots, x_r^\dagger) = V_e .$$

Then by considering lotteries with

$$(x_1^\dagger, x_{2*}^\dagger, x_3^\dagger, \dots, x_r^\dagger) \quad \text{and} \quad (x_1^\dagger, x_2^*, x_3^\dagger, \dots, x_r^\dagger)$$

as the basic reference consequences we can determine the shape of the u -curve between V_d and V_e . If the interval $[V_a, V_b]$ does not overlap with $[V_d, V_e]$, there remains a problem of matching up the u -scales between these two intervals. But surely we can be ingenious enough to get an overlap between $[V_a, V_b]$ and $[V_d, V_e]$, as in Figure 12, and we could use this overlap to straighten out the scaling problem. Still referring to Figure 12, we thus can extend the u -evaluation from the interval $[V_a, V_b]$ to the interval $[V_a, V_e]$. This might suffice for our purposes or else we would have to carry the extension process another stage.

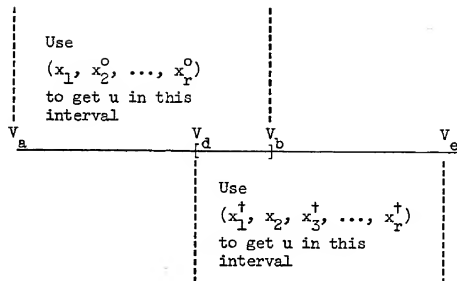


Figure 12

The processes of u -evaluation over a restricted domain and its extension to a wider domain should now be clear. There are lots of little details that need expansion but candidly more time has been spent on these ideas already than these ideas warrant either from a theoretical or a practical point of view.

We turn now to question #1: If Debreu-independence is satisfied, then how could we assess the v_i -functions in order to get the additive representation of V ? Without going into all the fine points, I will discuss below a procedure, which is based on the algebraic proof by Luce and Tukey [10] of the additive representation theorem.

In order not to get bogged down in a bunch of subscripts and superscripts I will confine myself to $r = 3$ and label the generic point in the space by (x, y, z) instead of (x_1, x_2, x_3) . I will eventually label points with subscripts such as $(x_{17.3}, y_{3.4}, z_{.8})$ and this point will be such that

$$V(x_{17.3}, y_{3.4}, z_{.8}) = 17.3 + 3.4 + .8 .$$

First choose a point as origin and label it (x_0, y_0, z_0) and let $V(x_0, y_0, z_0) = 0 + 0 + 0 = 0$. Next take an arbitrary unit in the x -direction and label this point (x_1, y_0, z_0) and let $V(x_1, y_0, z_0) = 1$. Let y_1 and z_1 be such that

$$(x_1, y_0, z_0) \sim (x_0, y_1, z_0) \sim (x_0, y_0, z_1) .$$

Next, let x_2 be the x -amount such that

$$(x_2, y_0, z) \sim (x_1, y_1, z) ,$$

that is, starting from (x_1, y_1, z) when we change the second coordinate from y_1 to y_0 we must compensate by changing the first coordinate from x_1 to x_2 , thus defining x_2 . Notice that by Debreu-independence

$$(x_2, y_0, z_0) \sim (x_1, y_1, z_0) \sim (x_1, y_0, z_1) .$$

We now proceed in an analogous manner to define y_2, z_2 , and then x_3, y_3, z_3 , and so forth. How about such things as $x_{.5}$? Suppose we tentatively take a

guess at $x_{.5}$. Then using this guess we can find $y_{.5}$ and $z_{.5}$ such that

$$(x_{.5}, y_0, z_0) \sim (x_0, y_{.5}, z_0) \sim (x_0, y_0, z_{.5}) .$$

To test whether this first approximation was reasonable, we could compare

$$(x_{.5}, y_{.5}, z) \quad \text{with} \quad (x_1, y_0, z) .$$

If the left alternative is less preferred, then we would choose a more pleasant x -value as our candidate for $x_{.5}$. We would iterate the procedure until we would get

$$(x_{.5}, y_{.5}, z) \sim (x_1, y_0, z) \sim (x_0, y_1, z) .$$

The formal justification of this step requires a continuity assumption. Now instead of trying to assess systematically such values as x_k for $k = .5, .25$, and so forth, at some stage we will be forced to use the natural order of the x -variable, plot the assessments we have made, and interpolate in-between values by "fairing" a smooth curve through the points previously assessed. This is done in Figure 13. We could look at the curve in Figure 13 in either of two ways: (1) for any real number k on the ordinate we could read off the x_k -value; or (2) for any x -value, where x is in its natural units, the corresponding ordinate is interpretable as $v_1(x)$. Similarly we would get $v_2(y)$, $v_3(z)$ and finally

$$V(x, y, z) \equiv v_1(x) + v_2(y) + v_3(z) .$$

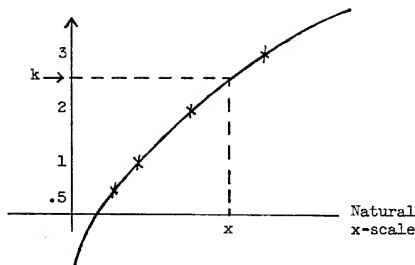


Figure 13

6.8. One More Proof of the Justification of Subjective Probabilities

As a corollary to Fishburn's additivity theorem we can get yet another justification of the introduction of subjective probabilities. To this end imagine that there are r "states of the world": $\theta_1, \theta_2, \dots, \theta_r$ and that one and only one of these states will prevail. Let the utility evaluation of any consequence C be denoted by $u_j(C)$ if state θ_j turns out to be true. Hence, not knowing at the time of decision which state is (or will be) true, any consequence C has an associated r -tuple $[u_1(C), \dots, u_r(C)]$. Still decisions have to be made and the DM must decide whether he prefers one r -tuple to any other r -tuple. Let's assume our DM wants to be consistent and wants to assign an overall utility function to these r -tuples.

In order to avoid fancy notation let us assume $r=3$, which introduces enough complexity to indicate how the general case would go. Let us suppress C and refer to the generic point as (u_1, u_2, u_3) . For each component consider two distinct u -levels: u_1' and u_1'' , u_2' and u_2'' , u_3' and u_3'' , and introduce four lotteries, each of which has two possible outcomes depending on the outcome of a fair toss. See Table 6.1 for details. Notice that for each lottery, if θ_j is true, there is a 50-50 chance of getting u_j' or u_j'' (for $j = 1, 2, 3$). In this context it seems eminently

TABLE 6.1

Toss	<u>Lottery</u>			
	1	2	3	4
H	(u_1', u_2', u_3')	(u_1', u_2', u_3'')	(u_1', u_2'', u_3')	(u_1', u_2'', u_3'')
T	(u_1'', u_2'', u_3'')	(u_1'', u_2'', u_3')	(u_1'', u_2', u_3'')	(u_1'', u_2', u_3')

reasonable, to me at least, that the DM should be indifferent amongst the four lotteries. But this is exactly the type of marginality assumption that is needed to prove the additive representation:

$$U(u_1, u_2, u_3) = \lambda_1 U_1(u_1) + \lambda_2 U_2(u_2) + \lambda_3 U_3(u_3) .$$

But since the components themselves are utilities, $U_j(u_j)$ is merely u_j and we have

$$U(u_1, u_2, u_3) = \lambda_1 u_1 + \lambda_2 u_2 + \lambda_3 u_3 .$$

I would now like to show that if the u_j 's are properly scaled, then the λ_j 's can be interpreted as the DM's judgmental probabilities of the states (i.e., of the θ_j 's). To this end, let us hypothesize the existence of several consequences $L_1, W_1, L_2, W_2, L_3, W_3$ which are defined as follows:

a. Let L_1, L_2, L_3 be such that the following stimuli are equally undesirable:

- $(L_1 \text{ with occurrence of } \theta_1) ,$
- $(L_2 \text{ with occurrence of } \theta_2) ,$
- $(L_3 \text{ with occurrence of } \theta_3) .$

Since these three stimuli are indifferent let us simply refer to any of these by L (mnemonic for "lose").

b. Let W_1, W_2, W_3 be such that the following stimuli are equally desirable:

- $(W_1 \text{ with occurrence of } \theta_1) ,$
- $(W_2 \text{ with occurrence of } \theta_2) ,$
- $(W_3 \text{ with occurrence of } \theta_3) .$

Since these three stimuli are indifferent let us simply refer to any of these by W (mnemonic for "win").

Let us scale u_j by letting u_j be 0 for L_j and be 1 for W_j . With this normalization the consequence $(1, 1, 1)$ is simply W and consequence $(0, 0, 0)$ is simply L . Hence $U(1, 1, 1) = 1$ and $U(0, 0, 0) = 0$. Now let's look at the meaning of λ_2 say. This quantity is such that the decision maker is indifferent between the stimulus $(0, 1, 0)$ and

between getting a λ_2 -chance at (1, 1, 1) -- i.e., at W -- and a $(1 - \lambda_2)$ chance at (0, 0, 0) -- i.e., at L. Hence the stimulus (0, 1, 0) means that the DM will get W if and only if θ_2 turns out to be true, and L otherwise. But the λ_2 value that equilibrates the options is, by definition, the judgmental probability that θ_2 will be true. A similar analysis holds for λ_1 and λ_3 .

7. Non-Additive Representations*

7.1. Quasi-Additive Representations

For the purposes of this section, think of the \underline{x} point as consisting of just two components, labelled (x, y) rather than (x_1, x_2) ; this labelling convention will avoid the excessive use of double subscripts. The reader should keep in mind, however, that the x and y components may themselves involve several components -- i.e., they may be vector quantities.

If u defined on (x, y) pairs has an additive representation, then

$$u(x, y) = u_X(x) + u_Y(y), \quad (7-1)$$

and from this we see that: (1) if the second component is held fixed at y_0 , then conditional (probabilistic) preferences on the x-component do not depend on the fixed value y_0 , i.e., in terms of the notation in Section 5.4, we have $X \text{ scui } Y$ -- (2) if the first component is held fixed at x_0 , then conditional (probabilistic) preferences on the y-component do not depend on the fixed value x_0 -- i.e., we have $Y \text{ scui } X$.

* The material in this section draws heavily on the research of Keeney [8, 9, 10].

In the sequel we shall also encounter the following representation:

$$u(x, y) = u_X(x) + u_Y(y) + \lambda u_X(x)u_Y(y) . \quad (7-2)$$

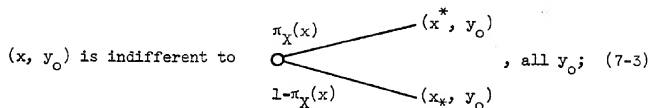
Now this representation also has the properties that $X \text{ scui } Y$ and $Y \text{ scui } X$ but if the constant $\lambda \neq 0$, then the representation is not additive. Let's examine closer the proposition $X \text{ scui } Y$. Hold the second component fixed at y_0 and consider u as a function of x . We get

$$\begin{aligned} u(x, y_0) &= u_X(x) + u_Y(y_0) + \lambda u_X(x)u_Y(y_0) \\ &= [1 + \lambda u_Y(y_0)]u_X(x) + u_Y(y_0) . \end{aligned}$$

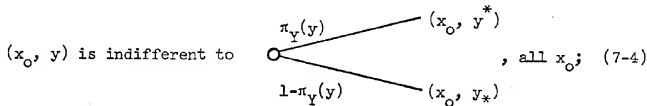
Hence we see that if, for example, (x_2, y_0) is indifferent to a lottery that gives an equal chance at (x_1, y_0) and (x_3, y_0) , this result would still obtain if the second component were changed from y_0 to any other common value. As far as gambles on the x -component are concerned, provided the second component is always kept at y_0 (say), the terms $[1 + \lambda u_Y(y_0)]$ and $u_Y(y_0)$ are constants, and, as such, do not essentially affect the choices on x -gambles.

In the remainder of this subsection I shall (1) show that if $X \text{ scui } Y$ and $Y \text{ scui } X$, then the representation given in (7-2) holds and (2) show how u_X , u_Y , and λ can be determined. In the next subsection we will discuss the case where λ can be taken equal to zero and therefore when an additive representation is appropriate. After that we will discuss the case where $X \text{ scui } Y$ holds but $Y \text{ scui } X$ does not.

Let x be constrained to lie between x_* and x^* and y between y_* and y^* .
Now since $X \underline{\text{scui}} Y$, we can define a function $\pi_X(x)$ such that



i.e., (x, y_0) is indifferent to a $\pi_X(x)$ -brlt with reference consequences (x_*, y_0^*) and (x^*, y_0) . Since $Y \underline{\text{scui}} X$, we can define a function $\pi_Y(y)$ such that



i.e., (x_0, y) is indifferent to a $\pi_Y(y)$ -brlt with reference consequences (x_0, y_*) and (x_0, y^*) . With these conventions we shall now prove that u has the representation

$$u(x, y) = a_1 \pi_X(x) + a_2 \pi_Y(y) + (1 - a_1 - a_2) \pi_X(x) \pi_Y(y) \quad (7-5a)$$

where

$$u(x_*, y_*) = 0, \quad (7-5b)$$

$$u(x^*, y_*) = a_1, \quad (7-5c)$$

$$u(x_*, y^*) = a_2, \quad (7-5d)$$

and

$$u(x^*, y^*) = 1. \quad (7-5e)$$

[Proof: From (7-3), (7-4) and the scaling and notational conventions of (7-5b, c, d, e) we have

$$\begin{aligned} u(x, y) &= \pi_X(x)u(x^*, y) + [1 - \pi_X(x)]u(x_*, y) \\ &= \pi_X(x)\{\pi_Y(y)u(x^*, y^*) + [1 - \pi_Y(y)]u(x^*, y_*)\} \\ &\quad + [1 - \pi_X(x)]\{\pi_Y(y)u(x_*, y^*) + [1 - \pi_Y(y)]u(x_*, y_*)\} \\ &= \pi_X(x)\{\pi_Y(y) + [1 - \pi_Y(y)]a_1\} \\ &\quad + [1 - \pi_X(x)]\{\pi_Y(y)a_2\} \\ &= a_1\pi_X(x) + a_2\pi_Y(y) + [1 - a_1 - a_2]\pi_X(x)\pi_Y(y). \end{aligned}$$

This is what we wished to show.]

The idea of the proof follows the pictorial display in Figure 7.1.

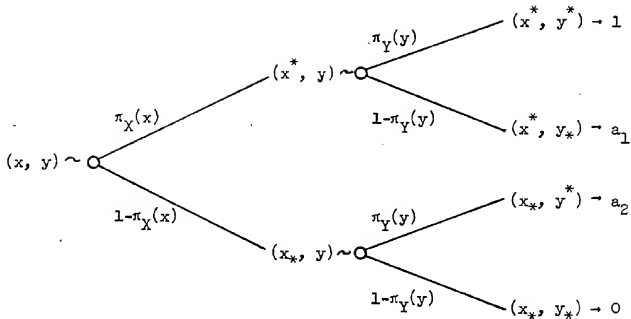


Figure 7.1

Thus we see that when we have both $X \text{ scui } Y$ and $Y \text{ scui } X$, then we can find the (joint) utility function, u , by getting:

- (1) the conditional utility function on x for given y_0 -- i.e., the function π_X ;
- (2) the conditional utility function on y for given x_0 -- i.e., the function π_Y ;
- (3) the value a_1 such that (x^*, y_*) for certain is indifferent to an a_1 -chance at (x^*, y^*) and (x_*, y_*) otherwise;
- (4) the value a_2 such that (x_*, y^*) for certain is indifferent to an a_2 -chance at (x^*, y^*) and (x_*, y_*) otherwise.

The representation of u in the form (7-2) or (7-5a) will be called a quasi-additive representation. If $\lambda = 0$ in (7-2) or $(1 - a_1 - a_2) = 0$ in (7-5a), then quasi-additivity becomes simple additivity.

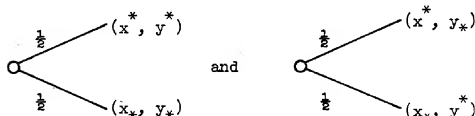
7.2. Quasi-Additive and Additive Representations

Let ℓ be a lottery that yields (x, y) pairs with given probabilities. We can think of ℓ as yielding a pair of random variables (\tilde{x}, \tilde{y}) which have a joint distribution. [The tilde notation is used to denote random variables.] If u is of the form (7-5a), then the expected utility of lottery ℓ is

$$Eu(\tilde{x}, \tilde{y}) = a_1 E\pi_X(\tilde{x}) + a_2 E\pi_Y(\tilde{y}) + (1 - a_1 - a_2) E[\pi_X(\tilde{x})\pi_Y(\tilde{y})] \quad (7-6)$$

Notice that $E\pi_X(\tilde{x})$ and $E\pi_Y(\tilde{y})$ depend respectively on the marginal distributions of \tilde{x} and \tilde{y} , whereas $E[\pi_X(\tilde{x})\pi_Y(\tilde{y})]$ depend on the joint distribution of \tilde{x} and \tilde{y} .

When can we assume the product term will disappear? The answer, of course, is when and only when $a_1 + a_2 = 1$; but this does not help very much. If, for example, the following two lotteries are indifferent



then we get, according to our scaling and notational conventions in (7-5b, c, d, e),

$$\frac{1}{2}(1) + \frac{1}{2}(0) = \frac{1}{2}(a_1) + \frac{1}{2}(a_2)$$

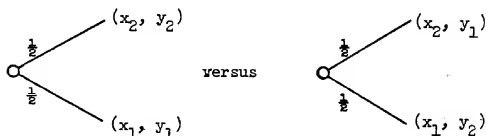
or

$$1 = a_1 + a_2$$

and hence

$$u(x, y) = a_1 \pi_X(x) + a_2 \pi_Y(y) .$$

It is not difficult to show that we do not have to use x_* , y_* , y^* , to test for additivity; any four values x_1 , x_2 , y_1 , y_2 where the test



yields indifference between the lotteries will suffice -- as long as $x_1 \neq x_2$ and $y_1 \neq y_2$.

But this test is a special case of the Fishburn-marginality assumption discussed in Section 6. Recall in that section we proved that if the marginality assumption held for x , x_* , y , y_* for all x , y pairs -- not just a single pair -- then the additivity representation holds and, a fortiori, we have $X \text{ scui } Y$ and $Y \text{ scui } X$. In this section we have shown that if $X \text{ scui } Y$, $Y \text{ scui } X$, and if the marginality assumption holds for any particular four components x_1 , x_2 , y_1 , y_2 (where $x_1 \neq x_2$ and $y_1 \neq y_2$), then additivity holds.

We can look at the Marginality Assumption as follows. Suppose a lottery l gives you a fifty-fifty chance at x_1 and x_2 and at y_1 and y_2 . Would you pay any positive premium for the privilege of pairing up x and y values? Would you prefer pairing x_1 with y_1 , or with y_2 ? If you would not be willing to pay any positive premium, then the marginality assumption holds.

Example 1. A manager has two separate businesses that he controls. Let x and y be the earnings per share for the two businesses respectively. Suppose that given y_0 the conditional utility function on x is $-e^{-\lambda_1 x}$, and given x_0 the conditional utility function on y is $-e^{-\lambda_2 y}$. In this case we have mutual scui. Now if the manager has a 50-50 chance at x_L or x_H and a 50-50 chance at y_L or y_H (where L and H are mnemonic for "low" and "high" respectively), he might prefer, if he were able, to pair y_H with x_L and pair y_L with x_H , so that the Fishburn Marginality Assumption would not hold. As a special case, the utility function might have the form

$$u(x, y) = -e^{-(\lambda_1 x + \lambda_2 y)}.$$

Example 2. A DM, upon retirement, might evaluate his post-retirement financial preferences in terms of an annuity income x and an amount y that he bequeaths to his survivors. He may feel that X and Y are mutually scui but given a 50-50 chance at x_L or x_H and another 50-50 chance at y_L or y_H he may prefer, if he were able, to pair y_L with x_L and pair y_H with x_H , so that the Fishburn Marginality assumption would not hold.

From the point of view of assessing utility functions the quasi-additive form (7-5) is not much harder to work with than the additive form, but the ensuing analysis is much harder. The quasi-additive form requires the full joint distribution of \tilde{x} and \tilde{y} and analytically we get involved with the complex term $E[u_X(\tilde{x})u_Y(\tilde{y})]$. Of course, if the random variables \tilde{x} and \tilde{y} are also probabilistically independent, then

$$E[u_X(\tilde{x})u_Y(\tilde{y})] = E u_X(\tilde{x}) \cdot E u_Y(\tilde{y})$$

and life is a bit simpler once again.

7.3. Use of Certainty Equivalents

Suppose u has the additive representation

$$u(x, y) = u_X(x) + u_Y(y) \quad (7-7)$$

where $u_X(x) \equiv a_1 \pi_X(x)$ and $u_Y(y) \equiv a_2 \pi_Y(y)$. Now consider the lottery ℓ whose outcome is the paired random variable (\tilde{x}, \tilde{y}) . Suppose \tilde{x} can take on the values x_1, \dots, x_m with probabilities p'_1, \dots, p'_m and \tilde{y} can take on the values y_1, \dots, y_n with probabilities p''_1, \dots, p''_n . Since u is of the form (7-7) we do not have to worry about the joint distribution of \tilde{x} and \tilde{y} . Now suppose \hat{x} is a certainty equivalent for \tilde{x} -- i.e.,

$$u_X(\hat{x}) = Eu_X(\tilde{x}) \equiv \sum_{i=1}^m p'_i u_X(x_i)$$

-- and \hat{y} is a certainty equivalent for \tilde{y} -- i.e.,

$$u_Y(\hat{y}) = Eu_Y(\tilde{y}) \equiv \sum_{j=1}^n p''_j u_Y(y_j).$$

In this case we can say the lottery ℓ with the random outcome (\tilde{x}, \tilde{y}) is indifferent to the certain outcome (\hat{x}, \hat{y}) , and this holds even though the random variables are jointly dependent in the probabilistic sense.

In other words, when the additive representation (7-7) holds it is legitimate to replace marginal distributions by certainty equivalents and we do not have to worry about probabilistic dependence.

Now let us investigate the case where the quasi-additive representation holds. If \tilde{x} and \tilde{y} are dependent random variables, then we cannot

say that l is indifferent to (\hat{x}, \hat{y}) where \hat{x} is the certainty equivalent of \tilde{x} and \hat{y} is the certainty equivalent of \tilde{y} . However, if \tilde{x} and \tilde{y} are independent random variables, then it is true that l is indifferent to the certain prospect of (\hat{x}, \hat{y}) . To see this let the probability of (x_i, y_j) be the p_{ij} where $\sum_j p_{ij} = p_i'$, $\sum_i p_{ij} = p_j''$, and where $p_{ij} = p_i' p_j''$. Now

$$\begin{aligned} Eu(\tilde{x}, \tilde{y}) &= \sum_{ij} p_{ij} u(x_i, y_j) \\ &= \sum_{ij} p_{ij} [a_1 \pi_X(x_i) + a_2 \pi_Y(y_j) + (1 - a_1 - a_2) \pi_X(x_i) \pi_Y(y_j)] \\ &= a_1 \pi_X(\hat{x}) + a_2 \pi_Y(\hat{y}) + (1 - a_1 - a_2) \sum_i p_i' \pi_X(x_i) \sum_j p_j'' \pi_Y(y_j) \\ &= a_1 \pi_X(\hat{x}) + a_2 \pi_Y(\hat{y}) + (1 - a_1 - a_2) \pi_X(\hat{x}) \pi_Y(\hat{y}) \\ &= u(\hat{x}, \hat{y}), \end{aligned}$$

as to be shown.

In summary: In order to use certainty equivalents \hat{x} for \tilde{x} and \hat{y} for \tilde{y} we need both

- a) mutual utility independence (more precisely, mutual scui), and
- b) probability independence of \tilde{x} and \tilde{y} .

If (a) is strengthened to require additivity of u , then (b) can be dropped.

7.4. Asymmetric Forms of Utility Independence

For notational convenience let us continue with a discussion of the case of two components x and y . Now let us suppose that we have X scui Y but not Y scui X . When we look at x -gambles we do not have to take into account which particular common value of y is under consideration but when we look at y -gambles the common value of x will be important. In

this case we can talk about a function $\pi_X(x)$ -- see (7-3) -- but we cannot define the function $\pi_Y(y)$. Knowing this, how can we build up the function $u(x, y)$? If \tilde{x} and \tilde{y} are independent random variables, can we make use of \hat{x} , the certainty equivalent of \tilde{x} ?

Keeney [8] shows that given the above assumptions, u can be determined if the following scaling conventions and assessments are made:

(1) let

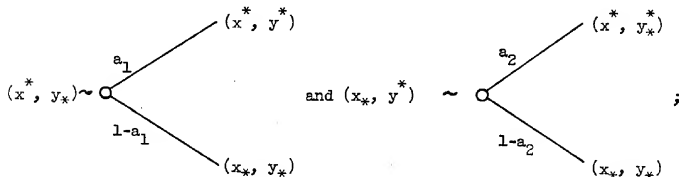
$$u(x^*, y^*) = 1,$$

$$u(x_*, y_*) = 0,$$

$$u(x^*, y_*) = a_1,$$

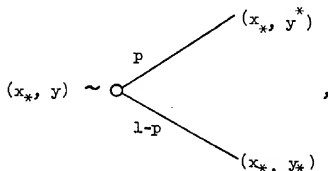
$$u(x_*, y^*) = a_2,$$

and determine a_1 and a_2 by the following indifferences:



(2) assess a function $\pi_X(x)$ as defined in (7-3);

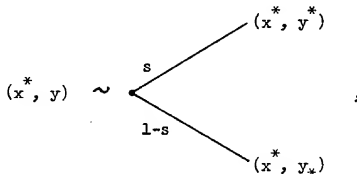
(3) assess a function $u(x_*, y)$ by looking at conditional preferences for y 's under the assumption that the first component is always x_* ; thus if



then p is functionally related to y by means of

$$\begin{aligned} u(x_*, y) &= pu(x_*, y^*) + (1 - p)u(x_*, y_*) \\ &= pa_2 ; \end{aligned}$$

(4) assess a function $u(x^*, y)$ by looking at conditional preferences for y 's under the assumption that the first component is always x^* ; thus if



then s is functionally related to y by means of

$$\begin{aligned} u(x^*, y) &= su(x^*, y^*) + (1 - s)u(x^*, y_*) \\ &= s + (1 - s)a_2 = a_2 + s(1 - a_2) . \end{aligned}$$

These assessments are shown in Figure 7.2.

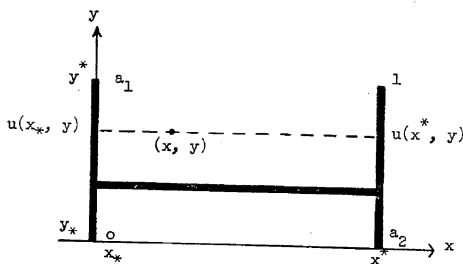


Figure 7.2

The heavy lines in the figure denote conditional utility assessments that have to be made. The heavy horizontal line can be shown at any y -level since $\pi_X(x)$ does not depend on the y -level. Now consider any point (x, y) where $x_* \leq x \leq x^*$ and $y_* \leq y \leq y^*$. We now have

$$u(x, y) = \pi_X(x)u(x^*, y) + [1 - \pi_X(x)]u(x_*, y) \quad (7-8)$$

and since we have assessed the functions $u(x^*, y)$ and $u(x_*, y)$, we can determine $u(x, y)$.

We have essentially shown that if we have $X \text{ scui } Y$ but not $Y \text{ scui } X$, then it is still not hopeless to determine u . In comparison to the quasi-additive form that required the assessment of constants a_1 , a_2 , and functions π_X , π_Y , we now need constants a_1 , a_2 , and functions π_X , $u(x_*, y)$, and $u(x^*, y)$.

Instead of assessing the two functions $u(x_*, y)$ and $u(x, y^*)$ we could get along with only one function $u(x_*, y)$ say and a single indifference curve in the (x, y) -plane provided this indifference curve ranges over all y -values from y_* to y^* . Without going into details let me sketch the argument. Suppose in Figure 7.3 the points on the wiggly line are all indifferent to each other. Two cases are considered.

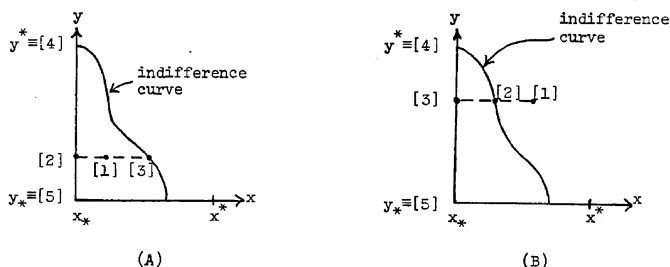


Figure 7.3

In Case (A) we consider any (x, y) point such as the one labelled [1]. Now [1] is indifferent to some π -chance at [3] and a complementary chance at [2], and π can be determined by the π_X -function. Let me write in a sloppy fashion

$$u[1] = \pi u[3] + (1 - \pi)u[2] .$$

But since [3] and [4] are indifferent, $u[3] = u[4]$. But now $u[2]$ can be determined from [4] and [5] using the $u(x_*, y)$ -function. Hence we see all that is necessary is: the $u(x_*, y)$ -function with any arbitrary

scaling, the indifference curve, and the $\pi_X(x)$ -function. If (x, y) is a point situated as [1] in Case (B), then we must express [2] as indifferent to a gamble between [1] and [3], viz

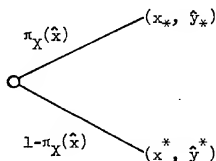
$$u[2] = u[1] + (1 - \pi)u[3]$$

where π comes from the π_X function. Now $u[2] = u[4]$ and we determine $u[3]$ by the $u(x_*, y)$ -function.

Now let's look at the use of certainty equivalents when X scui Y holds but Y scui X does not hold. Nothing can be done when \tilde{x} and \tilde{y} are dependent random variables. But now suppose \tilde{x} and \tilde{y} are independent random variables and let $p_{ij} = p_i^* p_j^*$ as is done in Section 7.3. In this case, suppose we let \hat{x} be the certainty equivalent of \tilde{x} , let \hat{y}_* be the certainty equivalent of \tilde{y} assuming $\tilde{x} = x_*$ and let \hat{y}^* be the certainty equivalent of \tilde{y} assuming $\tilde{x} = x^*$. Using the representation (7-8) we get

$$\begin{aligned} \sum_{i,j} u(x_i, y_j) p_{ij} &= \sum_{i,j} \{ \pi_X(x_i) u(x_*, y_j) + [1 - \pi_X(x_i)] u(x^*, y_j) \} p_i^* p_j^* \\ &= \pi_X(\hat{x}) u(x_*, \hat{y}_*) + [1 - \pi_X(\hat{x})] u(x^*, \hat{y}^*) \end{aligned}$$

Hence the lottery ℓ that yields the random outcome (\tilde{x}, \tilde{y}) is indifferent to the simple lottery



8. Value of a Life

8.1. Value of YOUR Life

In Section 5.5 we considered a problem in medical diagnostics-and-treatment and we showed how by successive reductions we could reduce a multi-attributed consequence of many dimensions down to a special pair of attributes: monetary considerations, x -- or perhaps instead, "pseudo-effective" number of days in bed with a specified level of discomfort -- and a probability p of an awesome occurrence such as death. In that section we deferred any consideration of the tradeoffs between x and p to the present section.

Many of the points I wish to make in this section can be illustrated by an ingenious example due to Richard Zeckhauser, who is delightfully gifted in being able to construct pithy, insightful, little brain-teasers that laybare methodological difficulties. Imagine that our protagonist is a merry bachelor who has no family ties, no special welfare projects he wishes to support, and accordingly who carries no insurance on his life. Through an intricate web of circumstances our protagonist finds himself as the unlucky player in a game of Russian Roulette. There are six chambers in a revolver and k of these chambers -- in the usual version, $k = 1$ -- are loaded; the rotating drum is spun in a way that each chamber has an equal chance of being used. Hence the bachelor has a $k/6$ chance of meeting his demise. We are interested in how much he should be willing to pay to remove just one bullet from the gun before the drum is spun? In particular, should he spend more to go from one bullet to none or from two bullets to one bullet?

Most people say he should pay more to go from one to none -- but they have second thoughts about their snap judgments when they think about going from six to five.

Let A denote the bachelor's asset position and let $\rho_k A$ be the amount he would be willing to spend to remove one of the k bullets from the drum. Let's look at what happens if $k = 1$. By the definition of ρ_1 he is indifferent between "buying a bullet," which would result in "life" with an asset position of $(1 - \rho_1)A$, or "not buying a bullet," which would result in $1/6$ -chance at death and a $5/6$ -chance at life with an asset position A . Remember that how much money the bachelor has after his death is of no importance to him. A decision diagram is shown in Figure 8.1.

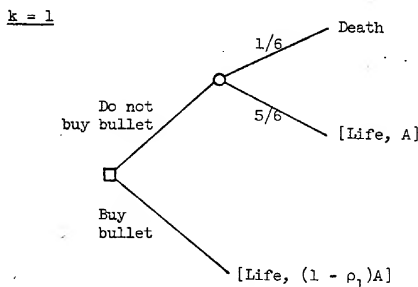


Figure 8.1

Notice that, if for example, $\rho_1 = \frac{1}{2}$, then we might be tempted to say, in a rather sloppy way, that the bachelor values his life at $3A$ since $1/6 \times 3A = \frac{1}{2}A$. This is just too sloppy a jargon to be of any use, since the "value of a life" we obtain in this fashion will depend on the probabilities involved. For example, the bachelor might be willing to pay

1/4 of his assets to remove a 1/100 chance at death and if we used the same principle, the bachelor would value his life at 25A. This is ridiculous!

Now let's turn to the case where $k = 2$. Figure 8.2 describes the strategic situation. The quantity ρ_2 is defined such that the two action

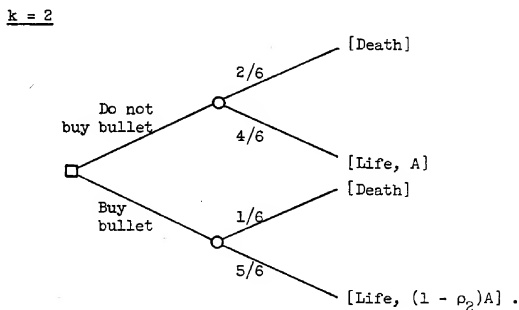
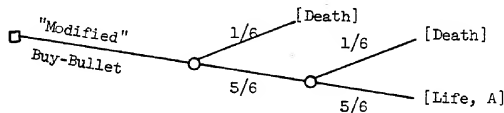


Figure 8.2

alternatives in Figure 8.2 are equally desirable. We shall now prove that consistency requires that $\rho_2 > \rho_1$. Suppose the converse were true, namely: $\rho_2 \leq \rho_1$ or equivalently $(1 - \rho_2) \geq (1 - \rho_1)$. In this case, if in Figure 8.2 we substituted the consequence [Life, $(1 - \rho_1)A$] for the consequence [Life, $(1 - \rho_2)A$], certainly the "Buy Bullet" alternative should not become more desirable. But we shall show that with this substitution

the Modified-Buy-Bullet alternative becomes more desirable and this contradiction demonstrates that $\rho_2 \leq \rho_1$ is untenable. By Figure 8.1 the consequence [Life, $(1 - \rho_1)A$] is just as desirable as a $1/6$ -chance at Death and a $5/6$ -chance at [Life, A]. Making these substitutions, we get for the Modified-Buy-Bullet path, the following:



which is reducible to a

$$\left(\frac{1}{6} + \frac{5}{6} \cdot \frac{1}{6}\right) \text{ or } \frac{11}{36}$$

chance at Death and a complementary chance at [Life, A]. But now we have a contradiction: assuming $(1 - \rho_1) \leq (1 - \rho_2)$, we substituted a worse (or at most an equally desirable) consequence in the Buy-Bullet branch of Figure 8.2 only to show that this transaction results in an $11/36$ chance at Death, and since the Do-Not-Buy-Bullet branch has a $12/36$ chance at Death, we see that the substitution of a worse consequence makes the resulting lottery better not worse. The trouble is we cannot assume $(1 - \rho_1) \leq (1 - \rho_2)$ or that $\rho_1 \geq \rho_2$. This proves that if the bachelor wants to be consistent, then $\rho_2 > \rho_1$, which means that he should pay more to reduce the chance at his death from $2/6$ to $1/6$ than to reduce this chance from $1/6$ to 0.

With the use of utility functions this same result can be proved as follows: Let $u[\text{Death}] = 0$ and $u[\text{Life}, A] = 1$. From Figure 8.1,

$$u[\text{Life}, (1 - \rho_1)A] = \frac{5}{6} .$$

From Figure 8.2, we have

$$\frac{4}{6} = \frac{5}{6} u[\text{Life}, (1 - \rho_2)A] ,$$

and therefore

$$u[\text{Life}, (1 - \rho_2)A] = \frac{4}{5} < \frac{5}{6} = u[\text{Life}, (1 - \rho_1)A] .$$

But if u is monotone increasing in assets (assuming "Life"), then this means

$$(1 - \rho_2)A < (1 - \rho_1)A$$

or

$$\rho_2 > \rho_1 ,$$

as we asserted.

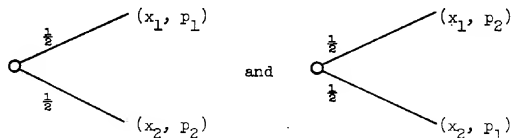
We leave as an exercise the assertion that consistency requires $\rho_1 < \rho_2 < \dots < \rho_6$.

Now let's see how our qualitative conclusions might change if we modify our assumptions.

(a) If our protagonist is not a bachelor but a family man, then the label "Death" may not be an adequate description of the outcome. He must worry about the financial state of his family when he contemplates Death and this must take into account any life insurance policies he may have.

(b) If the trigger of the revolver is not going to be pulled tomorrow but next year, then our analysis should not forget about his "anxiety along the way." In this case the bachelor might quite rationally -- that's a dangerous word to use -- pay more for the reduction of p from $1/6$ to 0 than from $2/6$ to $1/6$, since his anxiety during the year may, in some psychological sense, be almost as great with a probability of $1/6$ as it would be with $2/6$ and there is a world of difference between $1/6$ and 0. Hence the description of consequences must include the anxiety element since this may be of primary concern to the decision maker.

Let's return to the problem, posed in Section 5.5, where we wish to establish a utility function over pairs of attributes (x, p) , the first attribute x denoting income or assets and the second component p denoting the probability of Death. First let us see what we can do if we are willing to accept the rather strong Marginality Assumption, to wit: if a lottery gives a 50-50 chance at x_1 or x_2 and a 50-50 chance at p_1 or p_2 , then there is no incentive to match x_i and p_j pairs -- i.e., the decision maker is indifferent between the two lotteries:



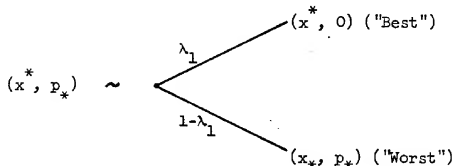
If this is true for all x_1, x_2, p_1, p_2 (in the range of our concern), then, as shown in Section 6.4, we have

$$u(x, p) = \lambda_1 u_1(x) + \lambda_2 u_2(p) . \quad (8-1)$$

Let the range of x be x_* to x^* and assume $0 \leq p \leq p_*$. Define

$$\begin{aligned} u(x_*, p_*) &= 0, & u(x^*, 0) &= 1, \\ u_1(x_*) &= u_2(p_*) = 0, & u_1(x^*) &= u_2(0) = 1. \end{aligned}$$

In this case $\lambda_1 + \lambda_2 = 1$, and in principle λ_1 is such that



It is very difficult to think about λ_1 using this characterization, and somehow we will have to think of other ways to get at this quantity.

Now let's consider the conditional utility function on p for a given x_0 value. Let the chance of Death, p , depend on whether a given event turns out to be S or F. If S, the chance at Death is p_1 ; if F the chance at Death is p_2 . See Figure 8.3. Let S and F be equally likely to occur. In this situation it is clear that the entire gamble boils down to a $(p_1 + p_2)/2$ chance at Death. In other words, by the very nature of the units on p it follows that $u_2(p)$ must be linear. But all is not so simple if one brings time into consideration. For suppose, in Figure 8.3, that event #1 will be resolved tomorrow and event #2 will be resolved next year. Then, starting from tomorrow on, the DM will know for the rest of the year that he has either a p_1 or p_2 chance at Death at the year's end. To be

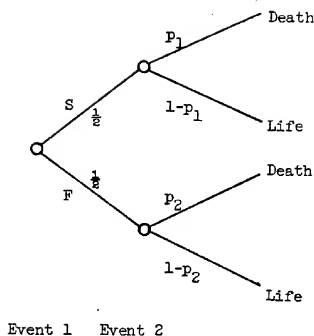


Figure 8.3

more concrete, suppose that the DM must choose between Option A: an outright .04 chance at Death (a year from now), and Option B: the composite gamble depicted in Figure 8.3 where $p_1 = 0.0$, $p_1 = .10$ and the two chance events are resolved tomorrow and next year respectively. With Option B he has a .05 chance at Death overall but tomorrow he will know whether $p = 0.0$ or $.10$. If he's lucky enough to get 0.0 , then he can avoid anxiety-along-the-way during the remainder of the year; if he's unlucky and ends up with $p = .10$, then he will suffer extreme anxiety-along-the-way but this may not be appreciably different than what he would suffer if he chose Option A. It might be worth it to him to choose Option B even though it might increase his objective chance at Death. There is no way of avoiding this psychological factor if the timing of events introduces this complication. In this case $u_2(p)$ will not be linear.

But suppose, to make things easy, that the resolution of event #2 in Figure 8.3 is immediate and there is no anxiety-along-the-way. In this case

$u_2(p)$ should be linear in p and, keeping in mind our normalization, we have

$$u_2(p) = 1 - \frac{p}{p_*}. \quad (8-2)$$

Substituting (8-2) into (8-1) we get

$$\begin{aligned} u(x, p) &= \lambda_1 u_1(x) + \lambda_2 \left(1 - \frac{p}{p_*}\right) \\ &= \lambda_1 [u_1(x) - \frac{\lambda_2}{\lambda_1 p_*} p] + \lambda_2 \end{aligned}$$

and since a utility function is meaningful only up to a positive linear transformation, we can express u in the form

$$u(x, p) = u_1(x) - \rho p \quad (8-3)$$

where $\rho > 0$. In other words, there is a constant substitution rate between p and x -utils.

In order to use representation (8-3) we must assess the critical parameter ρ . Before suggesting a way to do this, however, I would once again like to remind you of the role of sensitivity analysis. Thus, for example, in analyzing a choice between two lotteries involving (x, p) pairs it may be possible to compute a breakeven ρ , call it ρ_b , such that lottery 1 is preferred to lottery 2 if and only if $\rho \leq \rho_b$. In this case, the DM need only judge whether $\rho \leq \rho_b$.

How should we think about ρ ? If somehow we can find an (x_1, p_1) and an (x_2, p_2) pair that are indifferent in the eyes of the decision

maker, then we would have

$$u(x_1, p_1) = u(x_2, p_2) ,$$

or

$$u_1(x_1) - \rho p_1 = u_1(x_2) - \rho p_2 ,$$

and

$$\rho = \frac{u_1(x_1) - u_1(x_2)}{p_1 - p_2} . \quad (8-4)$$

Hence we see that to determine a $u(x, p)$ -function in the case where we are willing to assume the Marginality Assumption and where there is no anxiety-along-the-way, we merely have to assess a utility function for money (keeping in mind some constant probability of Death), and to find two (x, p) pairs that are indifferent to each other. It is not necessary to worry about the normalization of the utility function, u_1 .

If ℓ is a lottery with an uncertain payoff (\tilde{x}, \tilde{p}) , and if representation (8-3) holds, then regardless of the joint probabilistic dependence of random variables \tilde{x} and \tilde{p} we can use the certainty equivalents \hat{x} and \bar{p} , where \hat{x} is such that

$$u_1(\hat{x}) = Eu_1(\tilde{x})$$

and \bar{p} is merely the expected value of \tilde{p} .

It is easy to think of contexts (especially when the p -variable swings over a wide range) where the Marginality Assumption is untenable. However, even in this case we might still be willing to assume that X scui P (that is, preferences for lotteries involving x 's, with a common

value of p , do not depend on this common value of p) and $P \text{ scui } X$ (that is, preferences for lotteries involving p 's, with a common value of x , do not depend on this common value of x). In this situation an adaptation of the representation in (7-5) is appropriate. In that representation the second component was labelled y and was considered a desirable commodity, the more the better, not like p . It might be easier in our context if we change variables to let $y = 1 - p$ so that y is the probability of non-death. The rest of the adaptation is straight-forward; but, unfortunately, we must assess the two parameters a_1 and a_2 , which is not easy to do. Still, with only two parameters, it is possible to do some systematic sensitivity analysis.

Of the two assumptions $X \text{ scui } P$ and $P \text{ scui } X$, I feel that the latter is more usually tenable and most people, that I have asked, seem to agree with me. If we assume $P \text{ scui } X$ but not $X \text{ scui } P$, then the measurement problem is still more complicated, but not hopeless. We would need (according to Section 7.4) conditional utility functions on x for $p = 0$ and for $p = p_*$, as well as some normalizing constants.

Before going on to the next section let me add a few words about assessing some of these normalizing constants, like ρ in (8-3), like a_1 and a_2 in representation (7-5), and so on. Essentially, the decision maker is asked: Starting from a given point (x_1, p_1) how much would you be willing to give up in x to reduce p_1 to p_2 ? Think about this yourself: If $x_1 = \$100,000$ and $p_1 = .005$ how much would you be willing to give up to reduce p_1 to 0? Suppose you say \$90,000, so that for you

$$(\$100,000, .005) \sim (\$10,000, .00) .$$

In that case, if you had \$10,000 you would be willing to take a .005-chance at death for a payment at \$90,000. [If the numbers don't seem appropriate for you, add some zeroes where appropriate.] It's not easy to think of large amounts on one hand and small probabilities on the other hand and to corroborate any assessment it would be desirable to scrutinize this decision from different points of view. When this is done, it would be naive to expect consistency in the answers but by reflecting on the inconsistencies one might be able to better appreciate the issues at stake and come up with a more reflective compromise. One way to gain further perspectives into one's values about the value of life-saving is to employ a device suggested by Schelling [16] which involves chaining. Consider some undesirable state, that is not quite as bad as death, like blindness. Now ask such questions as:

(1) How much money would you be willing to give up to decrease the chance at blindness from p_3 to p_4 ?

(2) If you were certain to become blind and if you had the opportunity to undergo an operation which would cure you with probability $(1 - \pi)$ and kill you with probability π , how small would π have to be before you would be just willing to submit to surgery? We could now go on to insert deafness between life-as-usual and blindness. And so on. The trick should now be clear. By inserting a chain of undesirable consequences between Life and Death and by getting tradeoffs between certain consequences on one hand, and simple lotteries involving other awesome consequences on the other hand, we can by inference compute how we feel about giving up assets for decrements of the probability of death. Or, perhaps more realistically, we can get a

feeling for the order of magnitude for some of these constants and this may be all that is required for some decision-making purposes.

8.2. Value of a Statistical Life

Let us now switch our orientation and think about the decision problem facing a private or governmental agency which can spend additional amounts of money to reduce the proportion p of deaths in a given target population of size N . If $N = 100,000$ and at present $p = .0005$, so that 50 deaths will result this coming year from some discernible and partially preventable cause, should the agency spend an additional 10 million dollars to reduce p to $.0004$, a saving of 10 lives? It will probably make a difference if, after the fact, it is possible to unambiguously ascertain which of the fifty deaths could have been saved if the expenditures had been made. But should this matter? Would the agency be better advised to contribute the 10 million dollars to a group insurance plan, or to pay each member of the population \$100 for his own use?

There is a voluminous literature on this subject and I don't intend to summarize this literature here. Following the lead of Schelling [16], however, I would like to point out the connection of the contents of Section 8.1 with our present problem. Returning to the problem posed above, the agency administrator could, for example, ask each individual in the population how much he would be willing to forego in additional wages to save 10 lives in the population -- keeping in mind, of course, that he is part of the population. If the total amount, added over the 100,000 individuals exceeds \$10 million, then it is tempting to say that the agency would be better advised to spend the \$10 million on life-saving than on increased wages. But even this is not clear. A majority of the people might prefer

the added \$100 and there may not be any administrative device for monetary exchanges among the individuals. If the total amount, on the other hand, were less than \$10 million, it might still be advantageous for the agency to spend the money on life-saving. Notwithstanding these observations, it is, I believe, still relevant to inquire: how much is it worth to the target population? This is a point that Schelling makes quite forcefully in his paper.

There are two problems an agency administrator faces:

1. Given a budgetary amount for purposes of "life-saving" how should he distribute it amongst different projects? Naturally if his agency's past choices reveal that imputed tradeoffs between money and life are drastically different in various activities, then a realignment of effort and money might be called for. This is a problem of economic efficiency. Often (but not always) an analysis of this kind does not involve difficult and politically sensitive value judgments.

2. How large an overall budget should be spent on "life-saving"? This is a tougher question and is more in line with the subject matter of this paper. The tradeoffs between money (x) and a proportion (p) of anonymous deaths depends, as most things, on the particular environment of the problem. In some contexts the Marginality Assumption between x and p may be warranted; in others we might not be able to go that far but we might agree that $X \text{ scui } P$ or $P \text{ scui } X$. In some contexts we may get further simplifications because $u_1(x)$ or $u_2(p)$ may be linear. Let me close this section with a final word about the linearity of $u_2(p)$.

Suppose that in a given occupation there is a currently experienced proportion of .00005 of fatal accidents. To be dramatic, let us suppose an

innovation is suggested at zero cost that might double the accident rate with probability .3 and eliminate accidents entirely with probability .7. With the innovation the expected proportion is .00003 and this number is also the probability that a random individual from the population will suffer a fatal injury. Nevertheless, the agency administrator might prefer the currently experienced proportion of .00005 than the uncertain proportion with expected value of .00003. In other words the administrator's utility function defined on proportions may not necessarily be linear even though there is no question of anxiety-along-the-way for the individuals in the target population.

9. Intertemporal Tradeoffs

9.1. Introduction

An investor has to choose between two investment plans \underline{x} and \underline{y} . The stream of monetary payments and rewards he receives with \underline{x} is (x_1, x_2, \dots, x_m) , where x_i is his return in the i^{th} period; similarly, with \underline{y} he receives (y_1, y_2, \dots, y_n) . I use a different ending subscript, m for \underline{x} and n for \underline{y} , to emphasize the point that the streams can be of unequal length. Should he choose \underline{x} or \underline{y} ? If a return in a given period is positive he may use part of this return for consumption and part for further investment. The usual suggested procedure is to discount the stream, using some standard interest rate or cost of capital and the justification of the procedure is based on the observations: (1) money today can be used to make more money tomorrow, and (2) consumption today is sweeter than consumption tomorrow. If the x_i 's and y_j 's are uncertain (i.e., random variables), then several different techniques are suggested in practice: (1) discount the expected returns in each period, with possibly a higher discount rate to account for uncertainty, (2) discount certainty

equivalents at the certainty discount rate, (3) discount utilities at some rate, (4) discount the various possible certainty streams that could arise at the certainty rate, convert these present values to utilities and weight these by the respective probabilities of the streams.

In this section I would like to comment on these issues from the vantage point of the results developed in this paper.

9.2. The So-Called Certainty Case

It's not so easy in practice to find a "certainty case." For example, suppose that \underline{x} results in the stream (x_1, x_2, x_3, x_4) say where the four x 's are fully determined and where x_i is the return in the i th period. We now can engage in the procedure of successive reductions discussed earlier in Section 4. We could ask: If x_4 is reduced to zero, what compensating change must be made in x_3 to maintain indifference? In other words, we ask for a number x_3^* so that the following two streams are indifferent.

$$(x_1, x_2, x_3, x_4) \quad \text{and} \quad (x_1, x_2, x_3^*, 0) \quad . \quad (9-1)$$

Presumably x_3^* will depend on x_3 and x_4 ; but it also may depend on x_1 and x_2 . Suppose $x_4 > 0$ and $(x_3^* - x_3) > 0$. We could use some of the excess $(x_3^* - x_3)$ for investment in time 3 and some for consumption. Now even if the x_i 's are certain, we may not be certain about the investment opportunities at time 3. Furthermore, even if the two streams in (9-1) are indifferent it does not follow that the decision maker would follow the same consumption and investment choice in period 2 -- yes I mean 2! All these observations notwithstanding, we still can ask the investor to come up with some number x_3^* he is willing to live with, but we must recognize that this assessment might involve consideration of external factors in the environment that are themselves uncertain.

Next we seek a number x_2^* such that

$$(x_1, x_2, x_3^*, 0) \quad \text{and} \quad (x_1, x_2^*, 0, 0)$$

are indifferent. We repeat once again that x_2^* may depend on the possibly uncertain investment opportunities available in period 2. Finally we get an x_1^* such that

$$(x_1, x_2, 0, 0) \quad \text{and} \quad (x_1^*, 0, 0, 0)$$

are indifferent. By transitivity we now have indifference between

$$(x_1, x_2, x_3, x_4) \quad \text{and} \quad (x_1^*, 0, 0, 0) .$$

I will refer to x_1^* as the present value of the certain stream (x_1, x_2, x_3, x_4) .

There are some bounding relations that one should always keep in mind. For example, if there is available a standard investment that will take any amount Δ in period 3 and convert it into $\Delta(1 + \rho_3)$ in period 4, then, as one possible alternative, the investor might compare the streams in (9-1) with consumption held fixed in periods 1, 2, and 3 and from this conclude that

$$(x_3^* - x_3) \geq \frac{x_4}{(1 + \rho_3)} .$$

It is my impression that businessmen especially dislike standard discounting formulas, because these formulas do not sufficiently take into account the realities of their every day existence. "Sure I would

like a high net present value but not if it means that I won't get a good earnings statement this year and it is also critical that the increase of my earnings next year is better than my competitor's or else I will be in trouble with my board."

In handling long streams it sometimes is expedient to compromise by using a simple discounting procedure to discount a far-off future to a nearby future. For example, suppose the decision maker wishes to analyze the stream (y_1, \dots, y_{10}) ; he might decide to discount the payoffs y_4, y_5, \dots, y_{10} back to period 4 by letting

$$y_4^* = \sum_{j=4}^{10} \rho^{j-4} y_j$$

for some discount factor ρ and then to treat the streams

$$(y_1, \dots, y_{10}) \quad \text{and} \quad (y_1, y_2, y_3, y_4^*, 0, 0, \dots, 0)$$

as indifferent. However, he might be reluctant to use a similar discounting formula to bring the entire future back to the present. He might choose to be a bit more subtle in handling the immediate future.

9.3. Time Resolution of Uncertainty

The points I wish to make in this section can be illustrated by a simple example. Suppose we consider an investment whose payoff is an uncertain stream $(\tilde{x}_1, \tilde{x}_2, \tilde{x}_3)$ where the random variables are jointly dependent as shown in Figure 9.1. At time period 1, \tilde{x}_1 can take on the values -5 or -2 with probabilities .7 and .3 respectively. By -5 we mean that a net expenditure of five thousand (or million, if you like to think big) dollars has occurred. Suppose now in time period 1,

$\tilde{x}_1 = -5$; then in time period 2 there is a .4 probability of another loss of -4 and a .6 chance of 0 (i.e., no loss or gain). As a last illustration, if $\tilde{x}_1 = -2$, $\tilde{x}_2 = 0$, then there is a .9 chance that $\tilde{x}_3 = 2$.

In this single example there are eight possible streams ranging from (-5, -4, 2) to (-2, 8, 10) and these are listed together with their associated probabilities. For example the stream (-5, 0, 10) has a chance of occurring equal to $.7 \times .6 \times .5 = .21$.

Suppose now that we must choose between the uncertain stream $(\tilde{x}_1, \tilde{x}_2, \tilde{x}_3)$ and some other stream. We shall consider two versions of the problem.

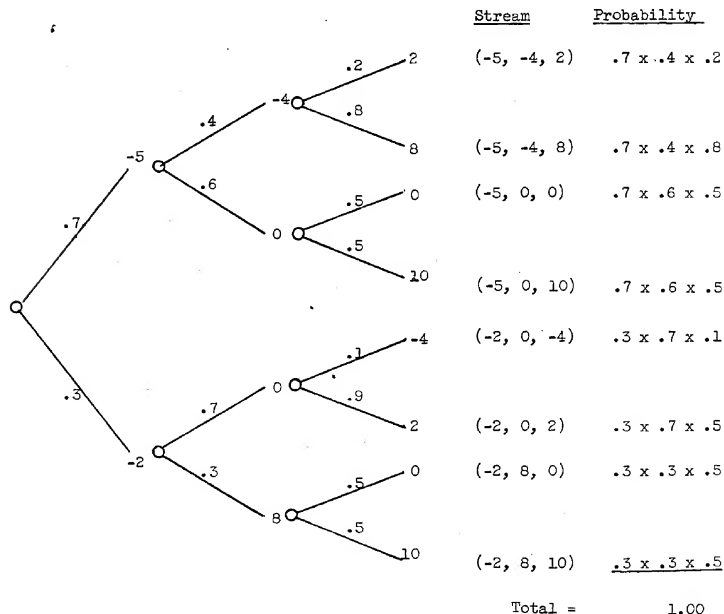


Figure 9.1

Version 1. If we choose \tilde{x} , imagine that immediately afterwards all uncertainties are resolved by means of some randomized device. The device will choose the stream $(-5, -4, 2)$ with probability $.7 \times .4 \times .2 = .14$, the stream $(-5, -4, 8)$ with probability $.7 \times .4 \times .8 = .224$, and so forth. If the randomized device turns up with $(-5, 0, 10)$, then the decision maker receives -5 in period 1, 0 in period 2, and 10 in period 3. In this version of the problem all uncertainties are resolved immediately.

Version 2. If we choose \tilde{x} , then at time period 1 we learn whether our payoff is -5 or -2 . At time period 2 we learn whether our payoff is -4 or 0 , (if $\tilde{x}_1 = -5$), or is 0 or 8 (if $\tilde{x}_1 = -2$). And so on. In other words, the uncertainty is resolved over time in a sequential manner.

In some cases there might be a tremendous difference between the two versions; in other cases the differences from a strategic point of view may be only slight, and even though the real problem might be as given in Version 2, it may be easier to analyze the problem by making believe it is posed as in Version 1. Sometimes, however, the analyst might inadvertently slip into using an evaluation procedure that is appropriate for Version 1 and not for 2. This is often the case where Monte Carlo techniques are employed. Imagine that the probability tree for \tilde{x} is more complicated: more components and more branches. The analyst might sample paths through the tree, evaluate each path as if it were a certainty, obtain a present-value equivalent for each path, and finally look at the empirical sampling (Monte Carlo) distribution of these present-value equivalents. This is fine for a Version 1 analysis but might yield misleading results for a Version 2 analysis.

Let us consider why these two versions are different. Suppose the investor is perched at time period two and compares two histories: $(x_1 = -5, x_2 = 0)$ and $(x_1 = -2, x_2 = 8)$. In both cases he looks forward to \tilde{x}_3 which can result in a fifty-fifty chance at 0 or 10. With history $(x_1 = -5, x_2 = 0)$ the investor may be in a state of deep anxiety. If perchance \tilde{x}_3 were to be zero, he might be in financial straights. Indeed he might at time 2 be willing to sell his option for \tilde{x}_3 for a certainty of 3 units. Whereas with history $(x_1 = -2, x_2 = 8)$ the investor might be quite relaxed and demand 4.5 units (say) for the uncertain \tilde{x}_3 option. With a history of $x_1 = -5$ the investor at time period 1 may not only be apprehensive about what might occur but he might take an action at that point in time that might insure against an ensuing catastrophe. But now you might say: why not bring these possible action alternatives into the formal description of the problem? We might be able to; but then again this might make the resulting problem so complicated that further analysis might bog down. When we treat a version 2 problem with a version 1 analysis, we may lose sight of anxiety-along-the-way and of the interaction between this particular problem and other problems in our environment. It is possible to compensate informally for these considerations. For example, suppose that a version 1 analysis produces by chance the certain stream $(-5, 0, 10)$. Now in evaluating this stream the investor might keep in mind the anxiety he would have suffered at periods 1 and 2 and also the fact that he might have altered some external plans at period 1. Accordingly, he might say that the certain stream of $(-5, 0, 10)$, keeping in mind these externalities, is equivalent (subjectively) to the "pure" stream $(-5, -2, 9)$ without these externalities.

An even simpler version of this problem might be of interest here.

A fair coin will be tossed: heads you will receive \$10,000 five years from now, and tails you receive nothing. How much would you want now for certain in lieu of this option? Well clearly it should matter to you if the coin is tossed tomorrow (immediate resolution) or tossed five years from now (delayed resolution).

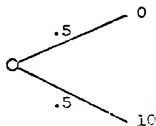
If the coin is tossed tomorrow, then tomorrow you will know whether you will have \$10,000 available five years from now. If, for example, you can invest \$8,000 tomorrow and be sure that this will yield \$10,000 in five years, then your present value of \$10,000 five years from now is worth at most \$8,000 tomorrow. For the sake of argument, let us suppose that you just as soon have \$7,000 tomorrow as \$10,000 five years from now. Hence today you face a lottery that will pay out \$7,000 or \$0 tomorrow with equal probabilities. Now your certainty equivalent for this gamble might be \$3,000, say. This could come from a direct assessment on your part, or, in more complicated situations, be computed from your utility curve for present assets. Alternatively, you might have asked yourself what amount certain would you want five years from now in lieu of the gamble -- still assuming the coin will be tossed tomorrow. You could then find your (subjective) present value of that certain amount. I personally find the latter approach less natural than the former approach and if my answers to the two approaches were to differ -- this would likely be the case -- I would feel more confident with the former approach.

If the coin were to be tossed five years from now, then in the next five years you will not know whether you can count on the \$10,000 or not. If you increase your consumption in anticipation of being lucky in the future, you will have anxiety-along-the-way. If you play it safe and you are lucky in the end, you will have lost the opportunity of happiness-along-the-way. Keeping

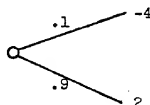
these things in mind, you might ask yourself what amount for certain would you take five years from now in lieu of the gamble. This is hard to think about because it involves a present perception of a future utility. You might, however, argue somewhat as follows: "If the gamble were to be conducted and payable tomorrow, I would have a certainty equivalent of \$4,000. But if everything were delayed five years my certainty equivalent would drop to \$3,000 (payable five years from now) because of the externalities we discussed above. Now a future payment of \$3,000, keeping in mind my investment opportunities and needs today, is only worth \$2,200 today, and this then is what I would want in lieu of the deferred gamble."

Notice that you might value the gamble at \$3,000 under version 1 (immediate resolution) and only at \$2,200 under version 2 (delayed resolution). [I was tempted at this point to write that this means you would be willing to pay \$200 to have the coin tossed tomorrow rather than five years from now, but that's not quite right. That error confuses buying and selling prices of uncertain contingencies.]

I think by this time it should be clear how we should analyze an uncertain x stream with immediate resolution (version 1). It might not be clear from this rambling discussion how we could analyze, at least in principle, a stream with delayed resolution. Please refer back to Figure 9.1. We could use a so-called roll-back or folding-back procedure. Suppose we consider ourselves at period two with a history of $(x_1 = -2, x_2 = 8)$, looking forward to the gamble, \tilde{x}_3 :



Let our present perception of our certainty equivalent for this future gamble be 4.5; hence we can replace the payoff (8 followed by the gamble) by the simple payoff $8 + 4.5$ or 12.5. Let the certainty equivalent of the gamble



from the vantage point of ($x_1 = -2$, $x_2 = 0$) be .5. Hence we now can replace the lower branch in Figure 9.1 by the reduced branch in Figure 9.2

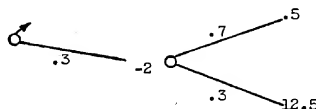
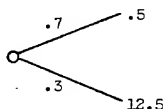
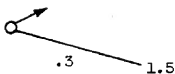


Figure 9.2

Now consider the gamble



from the vantage point $x_1 = -2$. Suppose its certainty equivalent is 3.5. Then we can replace the entire lower branch in Figure 9.1 by the simplified branch:



The procedure should now be clear (I hope!).

One obvious difficulty with using this procedure in practice is that it requires so many subjective assessments of certainty equivalents. In special cases, however, it may be possible to formalize these assessments by introducing utility functions. This will be the subject matter of a future paper.

In closing let me raise the additional issue of "present perception of a future utility." How can we account for changing tastes? Let \underline{x} be the summary of a typical consequence that is assigned at the tip of a decision tree and assume that this point is reached at some distant time period. We are now asked to evaluate \underline{x} using our present perception of our future tastes. But events not on the tree can influence these tastes. If these potential events are sufficiently important and determinable in advance we could list these as Q_1, Q_2, \dots, Q_r say; we could then assess the conditional utility of \underline{x} given Q_j —call it $u(\underline{x}|Q_j)$ —and finally compute

$$u(\underline{x}) = \sum_{j=1}^r u(\underline{x}|Q_j)P(Q_j) .$$

One might still feel, however, that there are other events not listed that could enter the picture and that the utility number $u(\underline{x})$ is really uncertain. But still we have to act, and the problem ultimately boils down to: If we have to make up our minds now, all things considered, would we rather have \underline{x} or \underline{y} ? By answering such questions we essentially force ourselves to assess a certainty utility equivalent, $\widehat{u(\underline{x})}$, for this uncertain $\widetilde{u(\underline{x})}$ and this $\widehat{u(\underline{x})}$ is interpreted as our utility function as of this time that is relevant for making decisions now.

This all implies the need for further work on preferences for consumption streams. It also provides, if not an appropriate, at least an essential stopping point for this first phase of a continuing research undertaking. In the future, I plan to continue my research on a series of closely related topics in this area grouped generally under the following headings:

- (a) preferences for multi-attributed alternatives with particular emphasis on benefit or consumption streams,
- (b) techniques for assessments of judgmental probability distributions,
- (c) assessments of preferences and judgments by panels of experts, and
- (d) use of heuristics and incomplete trees to structure very complex decision problems.

BIBLIOGRAPHY

1. Debreu, G., "Topological Methods in Cardinal Utility Theory," in K.J. Arrow, S. Karlin, and P. Suppes (eds.), Mathematical Methods in the Social Sciences, 1959, Stanford University Press, Stanford, California, 1960.
2. Fishburn, P.C., "Independence in Utility Theory with Whole Product Sets," Operations Research, Vol. 13, 1965, pp. 28-45.
3. _____, "Additivity in Utility Theory with Denumerable Product Sets," Econometrica, Vol. 34, 1964, pp. 500-503.
4. _____, "Additive Utilities with Incomplete Product Sets: Application to Priorities and Assignments," Operations Research, Vol. 15, 1967, pp. 537-542.
5. _____, "Interdependence and Additivity in Multivariate, Unidimensional Expected Utility Theory," International Economic Review, Vol. 8, 1967, pp. 335-342.
6. _____, "Additive Utilities with Finite Sets: Applications in the Management Sciences," Naval Research Logistics Quarterly, Vol. 14, 1967, pp. 1-13.
7. Hammond, J.S., "Bringing Order into the Selection of a College," Personnel and Guidance Journal, March, 1965, pp. 654-660.
8. Keeney, R.L., "Decision Analysis and Utility Assessment for Multidimensional Consequences," TIMS XV International Meeting, Cleveland, Ohio, Sept. 11-13, 1968.
9. _____, "Evaluating Multidimensional Situations Using a Quasi-Separable Utility Function," IEEE Transactions on Man-Machine Systems, Vol. MMS-9, June, 1968, pp. 25-28.
10. _____, "Quasi-Separable Utility Functions," Naval Research Logistics Quarterly, (to appear, December, 1968).
11. Krantz, D.H., "Conjoint Measurement: The Luce-Tukey Axiomatization and Some Extensions," Journal of Mathematical Psychology, Vol. 1, 1964, pp. 248-277.
12. Luce, R.D. and Tukey, J.W., "Simultaneous Conjoint Measurement," Journal of Mathematical Psychology, Vol. 1, 1964, pp. 1-27.
13. Manheim, M.L. and Hall, F.L., "Abstract Representation of Goals," P-67-24, Department of Civil Engineering, M.I.T., January, 1968.

14. Miller, J.R., III, "The Assessment of Worth: A Systematic Procedure and Its Experimental Validation," Doctoral Dissertation, M.I.T., June, 1966.
15. Raiffa, Howard, Decision Analysis: Introductory Lectures on Choices under Uncertainty, Addison-Wesley, 1968.
16. Schelling, Thomas, "The Life You Save May Be Your Own," in Samuel B. Chase, Jr., (Ed.), Problems in Public Expenditures Analysis, Brookings Institution, 1968.

